Distributed Gauging Methodologies for Variation Reduction in the Automotive Body Shop

by

Robert Eugene York

B.S. Mechanical Engineering
Purdue University, 1993

Submitted to the MIT Department of Mechanical Engineering and to the Sloan School of Management in partial fulfillment of the requirements for the Degrees of

Masters of Science in Mechanical Engineering
and
Masters of Science in Management

at the

Massachusetts Institute of Technology
June, 1995

Signature of Author ______________________________
Department of Mechanical Engineering
Sloan School of Management
May 12, 1995

Certified by ______________________________
David E. Hardt
Professor of Mechanical Engineering, Acting Co-director LFM

Certified by ______________________________
Roy E. Welsch
Professor of Statistics and Management Science

Accepted by ______________________________
Ain A. Sonin
Chairman, Committee on Graduate Students, Mechanical Engineering

Distributed Gauging Methodologies for Variation

Reduction in the Automotive Body Shop

by

Robert Eugene York

Submitted to the MIT Department of Mechanical Engineering and to the Sloan School of Management in partial fulfillment of the requirements for the Degrees of

Masters of Science in Mechanical Engineering
and
Masters of Science in Management

Abstract

This paper is concerned with methodologies to help in variation reduction of the body in white (BIW) in automotive body shops. In the production of the BIW, as with any manufactured product, variation is a concern. The main concern with variation is that variation is expensive for the company. The trick is to find which sources of variation can be eliminated, reduced, or counteracted.

For the most part, industrial applications have failed to keep up with advances in hardware and data collection. The data for this research was collected from in-line Optical Coordinate Measuring Machines at a General Motor's assembly plant. These stations were capable of measuring 100% of the major sub-assemblies and BIWs.

Because of a shortage of statistically trained analysts and a lack of tracking, the data are often only examined on a station by station basis, and then often via an assortment of linear univariate methods. There is a critical need, and much opportunity, to discover relationships via linear and non-linear multivariate analysis of data across many stations.

It was quickly discovered, however, that the necessary "building blocks" for a study of that kind were missing. Therefore, this research established methodologies that could be used to prepare data for studying the effects that upstream inputs can have on the downstream body in white (BIW). Four methodologies are presented: tracking parts, removing outliers, identifying presentation error, and evaluating the BIW results. These methodologies are essential in conducting variation reduction over a distributed gauging system. While these methodologies have been applied to data from an automotive shop, they should be applicable in other manufacturing settings as well.

Thesis Advisors:
David E. Hardt, Professor of Mechanical Engineering and Acting Co-director LFM
Roy E. Welsch, Professor of Statistics and Management Science

Acknowledgments

First, and most importantly, I wish to thank my wife, Lisa, for her unselfish support and sacrifices these past two years.

I would not be where I am without the desire for learning, which was instilled by my parents, to whom I am eternally grateful.

The Technical Center and an assembly plant of General Motors deserve special mention as hosts of my internship. I would like to thank them for extending their hospitality and providing support and resources during the course of my internship.

Many people aided me during the course of my internship. It would not be possible to enumerate everyone, but I would like to give special recognition to a few at GM. Harlan Neuville served as my industrial supervisor throughout this thesis. His support, as well as the support from his group, helped to make the project a possibility. James Clinton provided indispensable help and expertise in the assembly plant. Perhaps, most of all, I am indebted to Marshall Galpern. Marshall has devoted a lot of time and effort on my behalf over the past year. His insights are very much a part of this work. To all of them and everyone else at GM, I am very grateful.

My advisors, David Hardt and Roy Welsch, were always available with suggestions and encouragement whenever I sought their counsel. They certainly contributed to my learning throughout the whole experience and I would like to thank them for all their help. I also enjoyed working with Vikas Sharma, a fellow MIT student. Vikas was kind enough to help by sharing his mind in times of need.

This material is based upon work supported under a National Science Foundation Graduate Research Fellowship. Any opinions, findings, conclusions or recommendations expressed in this publication are those of the author and do not necessarily reflect the views of the National Science Foundation.

The research in this thesis was conducted as part of an internship, made possible under the auspices of the Leaders for Manufacturing Program, a partnership between MIT and major US manufacturing corporations. I am deeply grateful to the LFM program for the internship opportunity and for the resources and support I received these past two years as a Fellow.

Table of Contents

Chapter 1 - Introduction

This paper is concerned with methodologies to help in variation reduction of the body in white (BIW) in automotive body shops. For those not familiar with the auto industry, chapter two will describe what composes a BIW. But for now the production of the BIW, as with any manufactured product, is concerned with variation. While these methodologies have been applied to data from an automotive shop, they should be applicable in other manufacturing settings as well.

Some sources of variation are known. They include changes in the raw material inputs, changes in the process, and changes in external factors, just to name a few. The trick is to find which sources can be eliminated, reduced, or counteracted. The other alternative is to make the process robust to the sources.

The main concern with variation is that variation is expensive for the company. High variation can increase scrap and rework costs. Variation affects part interchangeability, which pushes up costs because the assembly process must accommodate parts with different dimensions. Variation is related to customer satisfaction with things such as fit and finish of the final product. Warranty costs are also impacted by variation, which can increase the chance that the automobile will have leaks.

It was initially intended that this research would be a multivariate analysis across a distributed gauging system. It was quickly discovered, however, that the necessary "building blocks" for a study of that kind were missing. Therefore, this research sets up methodologies, depticted in Figure 1, that can be used to prepare data for studying the effects that upstream inputs can have on the downstream BIW.

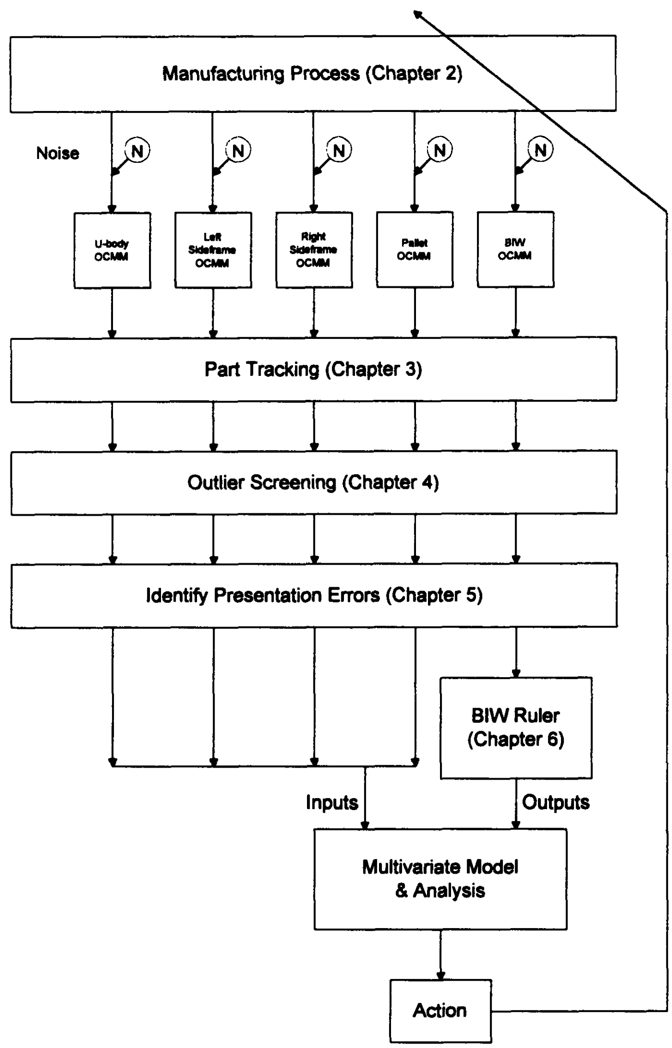


Figure 1 - These methodologies prepare data for a multivariate analysis

After covering the background information, each chapter addresses a methodology that was used on data collected from a body shop. These methods include establishing part tracking, handling outliers, identifying presentation error, and judging the output. The final chapter summarizes the conclusions of the research.

Each of the methodology chapters is structured to make them somewhat autonomous. They are structured in this way to facilitate not only the transfer of individual methodologies to other industries, but also for the readers who are interested in just one methodology and do not wish to read the entire thesis. Each chapter answers 1) what is the methodology, 2) why is it important, 3) how is the problem currently handled, 4) how was this methodology implemented, and 5) what are the conclusions for this methodology.

Chapter 2 - Background

The purpose of this chapter is to provide just enough background on the automotive body shop and the Optical Coordinate Measuring Machines (OCMM). This chapter will provide enough information to understand the remaining chapters and will provide references for those wishing to study these subjects more thoroughly.

Automotive Body Shop

The automotive body shop is responsible for building the sheet metal structure of the automobile. A major milestone in the process is the production of the body in white (BIW). The BIW, Figure 2, is the full sheet metal structure before anything is hung on or bolted to it. It is comprised of roughly three hundred individual parts welded together.

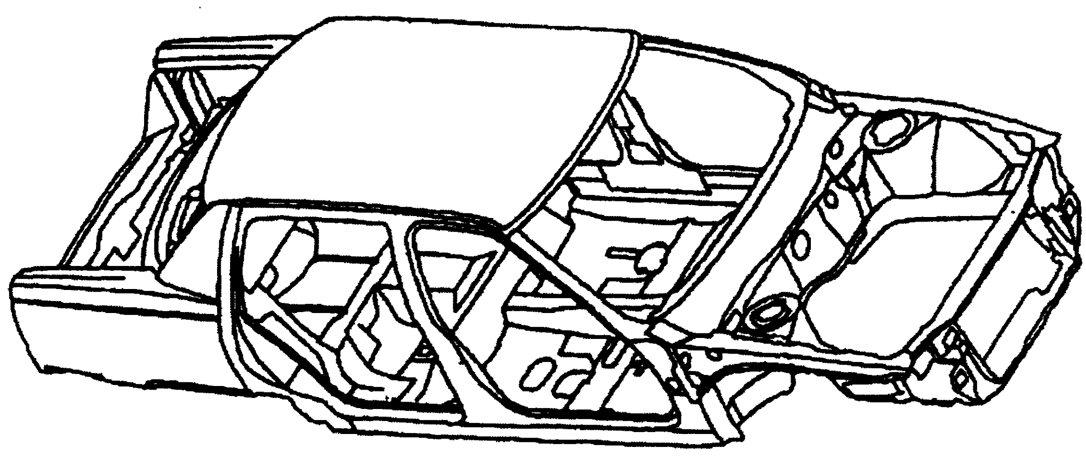


Figure 2 - The body in white is composed of around 300 welded pieces of sheet metal

The process of putting those three hundred parts together is simplified by the production of sub-assemblies. Three major sub-assemblies, Figure 3, along with the roof comprise the BIW. These three sub-assemblies are the left sideframe, the right sideframe, and the underbody.

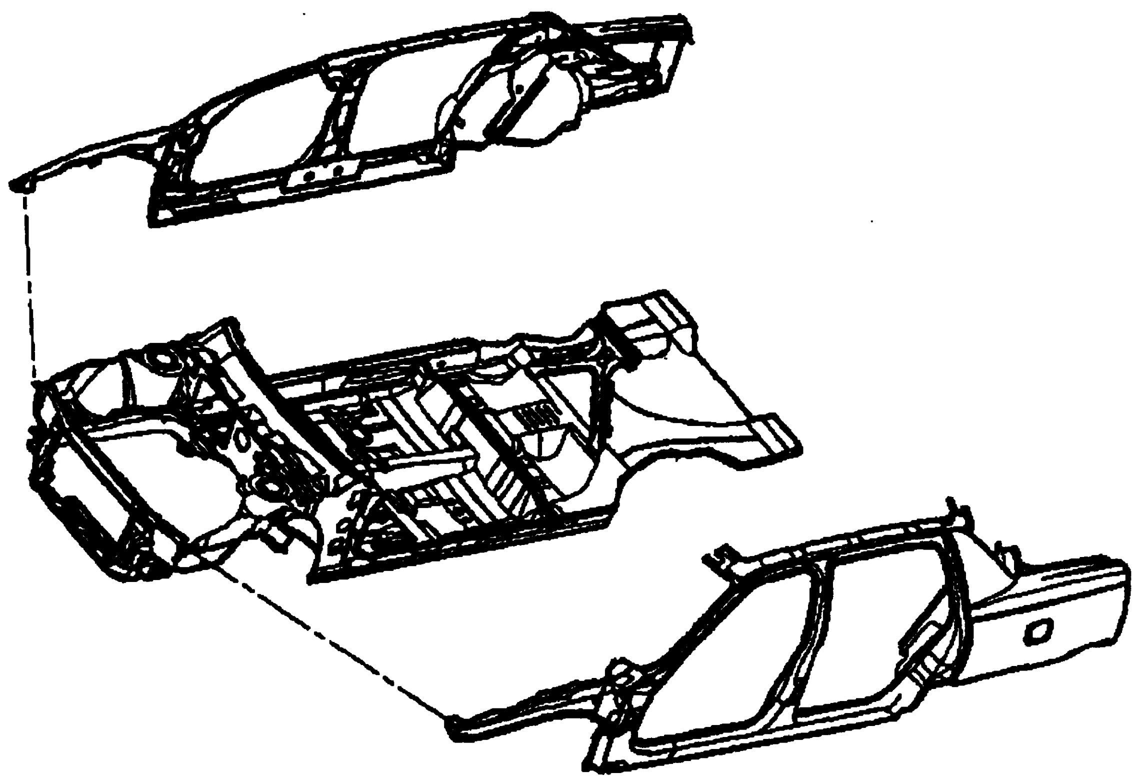


Figure 3 - The underbody and left and right sideframes are the BIW's major subassemblies

At the GM assembly plant, these three major sub-assemblies are married together on a pallet. A pallet is the tooling that locates and fixtures the underbody with pins and pads located on the pallet. A number of pallets circulate around in a loop, stopping at various stations as the sideframes and roof are welded to the underbody. When the BIW is finished, the pallet returns on the loop to pick up a new set of sub-assemblies.

The primary function in the body shop is the welding of sheet metal parts. The main goal of the processes is to produce consistent, quality parts. The environment, weld parameters, initial fitup, fixturing parameters, and joint design are just a few of the factors that affect the final quality of the BIW (Pool, 1991). Figure 4 is an Ishikawa (fishbone) diagram that graphically depicts some of the sources of variation in the BIW.

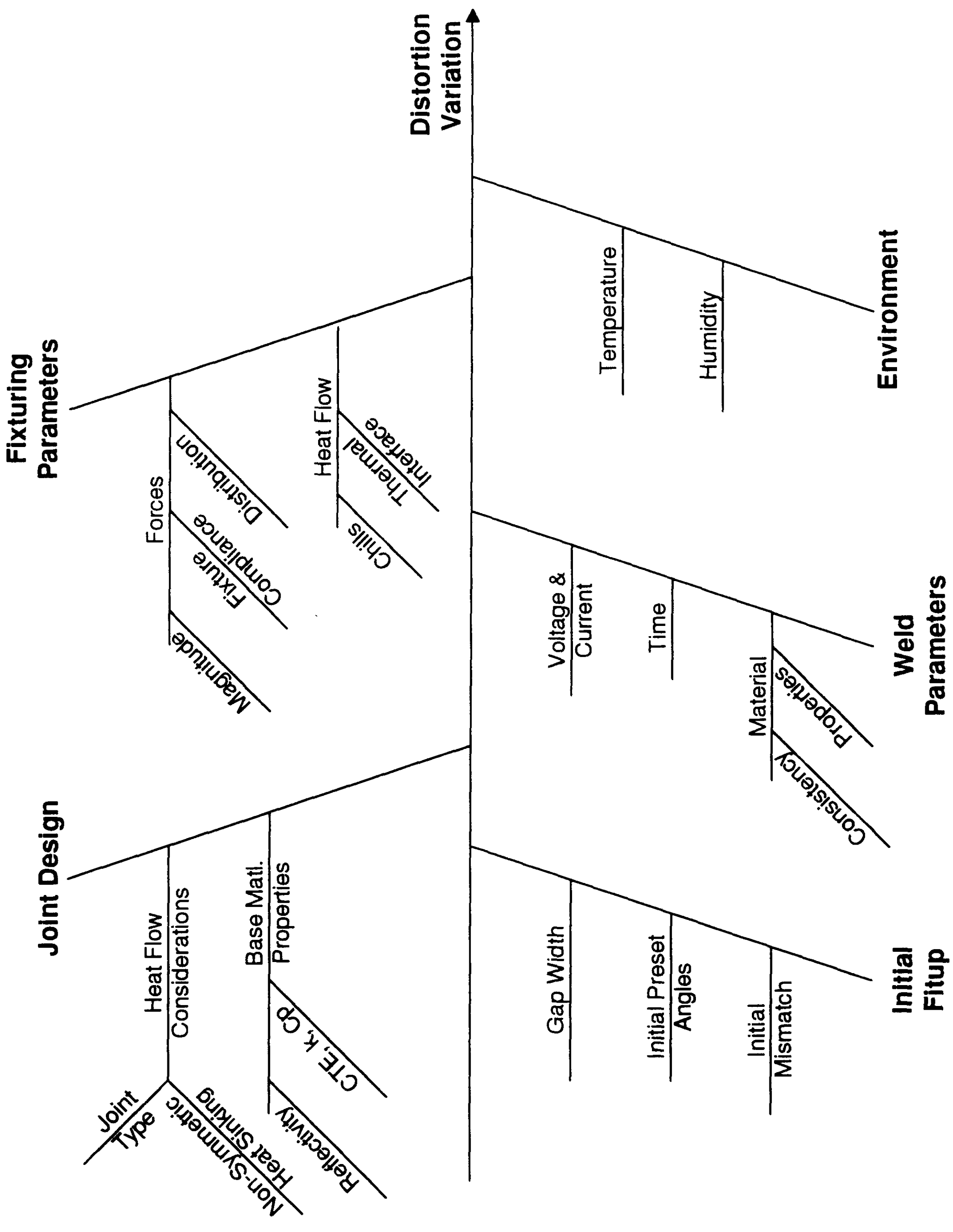


Figure 4 - An Ishikawa diagram highlights some of the sources of BIW variation

The body shop processes that produce the BIW are a small part of the automobile production line. The sheet metal parts that become the sub-assemblies begin as rolls of sheet metal that are stamped in presses at a stamping facility. Through a series of stations, the parts become sub-assemblies that fit into larger sub-assemblies that eventually end up becoming part of a BIW. The BIW travels to other sections of the body shop where doors, a hood, and other sheet metal components are added. The paint shop accepts the automobile body from the body shop and prepares it for general assembly. General assembly adds the remaining parts and touches to the automobile and prepares the car for shipping.

For a detailed look at BIW processes and broader view of the automobile production process, consider reading Jay Baron's thesis on dimensional analysis and process control of BIW processes (Baron, 1992).

Optical Coordinate Measuring Machines

The majority of data used in this research were collected using OCMM stations installed in the body shop. Each measurement station consists of a controller and several "cameras." Each camera is responsible for measuring one feature on the part or tooling. The cameras straddle the production line, Figure 5, taking non-contact measurements on each part that stops in the station.

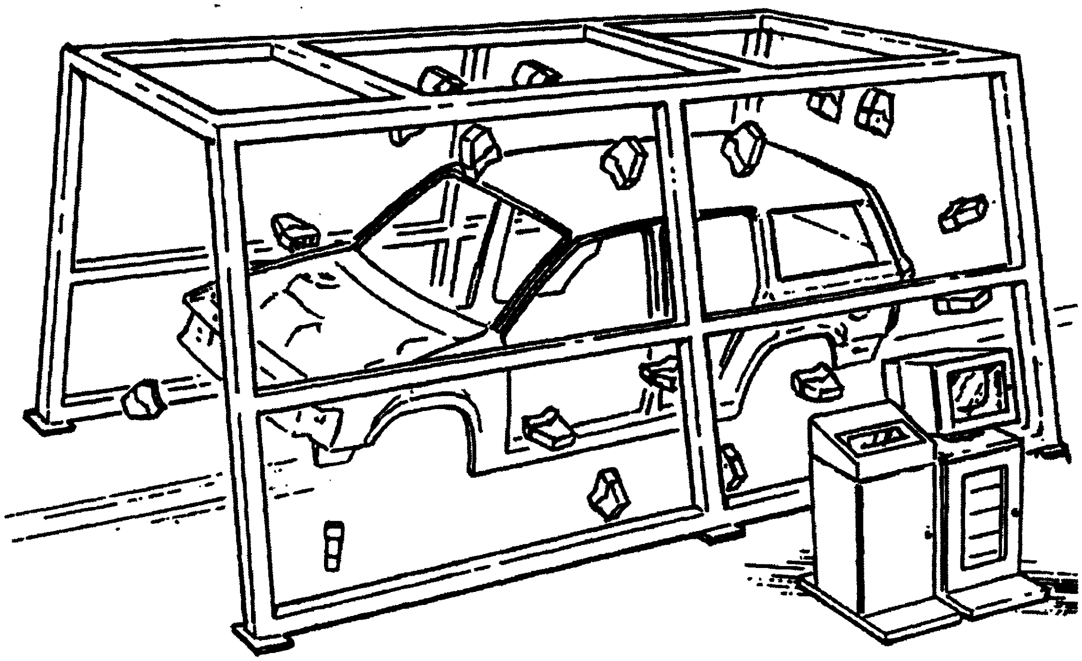


Figure 5 - An OCMM measurement station measures parts on the production line

The cameras emit a plane of laser light onto the part and use structured light sensors within the camera to obtain an image of the feature (Dewar, 1994). The data from each camera is passed to the controller. The controller analyzes the data with an algorithm appropriate to the feature being measured (Perceptron, 1993 and 1994). The execution of the algorithm results in either a measurement or an error code. The measurements from the various cameras can be converted into one reference frame. The error codes are used to help diagnose measurement problems.

The GM assembly plant contains eight OCMM stations. The stations are located in the body shop and measure:

- BIW - all three models
- Pallet - identical for all models
- Left Sideframe - one model
- Right Sideframe - one model
- Left Sideframe - two models
- Right Sideframe - two models
- Underbody - all three models
- Floorpan - all three models

The OCMM systems appear to have some advantages over manual and CMM measurement techniques. Not only do OCMMs require a third of the manpower of conventional techniques, the accuracy of the measurement is at least as good (and usually more precise) than other systems (Pizzimenti, 1994). The most significant gains are found in the speed and coverage of the measurements. The OCMM stations are imbedded in the production line and are capable of measuring every part. Manual and touch probe techniques take longer for a set of measurements and cannot keep pace with the production line. Table 1 replicates Pizzimenti's comparison of the three typical body shop measurement techniques.

Table 1 - The OCMM excels in speed, accuracy, and flexibility

	Manual	**CMM**	**OCMM**
Method	Contact	Contact	Non-Contact
Accuracy	0.5 mm	0.1 mm	0.1 mm
Manpower	3	3	1
Parts Measured	1/1000	1/250	1/1
Time	8 hours	4 hours	15 seconds
Cost	$200,000	$500,000	$250,000
Reconfigurability	No	Yes	Yes

Chapter 3 - Tracking Parts

What is Tracking?

To answer the question "what is tracking," consider an analogy to the food industry. The food industry incorporates tracking on its products. Each can of soup, for example, is labeled with a lot number. That lot number is the key to the entire history of that can of soup. If the can were found to be contaminated, they could use the lot number to find out:

- The time and place the soup was canned
- When and where the cans originated
- The lots and sources of the food in the can
- Where the rest of the lot was shipped

The food industry incorporates tracking because they need to be accountable and responsive when a problem is found, but why is tracking not being used in an automobile body shop?

Perhaps, the legacy of mass-production has kept tracking out of the plants. Mass production was built on the theory of interchangeable parts. Since each part was "interchangeable," there was no need to identify each part individually. However, with an increase in consumer awareness of quality, the automobile industry has had to mass produce to tighter tolerances. The increased difficulty in meeting these new tolerances uncovers the reality that each part is indeed different.

The costs of implementing tracking in the body shop outweighed the needs for tracking until now.

Why is Tracking Important?

Tracking releases the constraint of group or aggregate studies and allows direct analysis of the upstream measurements to the downstream measurements. The one-to-one match increases the sensitivity and confidence of the analysis.

Furthermore, automatic and electronic tracking provides large data sets of traced measurements. These large data sets are needed in some analysis methods that could not be performed if the current tracking methodology were being used. For example, traditional multivariate statistical methods such as principal components analyses require more samples (automobiles) than there are measured parameters (typically hundreds).

How is Tracking Currently Handled?

Tracking in the automotive body shop is usually done either automatically on every part or manually for small studies. Table 2 indicates the frequent methods that are implemented to accomplish tracking.

Table 2 - Automatic and manual tracking methodologies are used in body shops

	Automatic	Manual
Information Systems	X	
Radio Frequency ID	X	
Barcoding	X	X
Hand Marking		X

Information systems capitalize on existing material handling and scheduling systems to track parts. While this method does not require any contact with the parts, it must be robust to exceptions. An exception would occur, for example, if one part is taken off the line and another put in its place.

Radio frequency identification (RFI) is a highly flexible solution that overcomes the limitation of exceptions. A RFI system places a passive tag on the automotive part. Because of this, a RFI system has material costs, equipment costs, and labor costs (usually maintenance). RF tags can carry product identification, specific instructions, and other data for automated operations.

It is the transceiver, which is mounted near the production line, that provides the energy for data transmission and reception between the tag and transceiver. An electromagnetic field generated by the transceiver determines the dimensions of the transmission zone. As a tag enters the transmission zone, data transfer takes place without contact. RFI systems are usually integrated with the information systems.

Barcoding places a bar-code on the automotive part. For automatic barcoding, the bar-code is painted directly on the part. The bar-code is removed during the regular cleaning operation at the start of paint operations. During manual studies, the bar-code may be applied using stickers, which must be manually removed prior to paint operations. The bar-code is read using laser or light scanners, which are also typically integrated with the information systems.

The manual method, illustrated in Figure 6, is easily implemented for small studies. The method involves writing a number or putting a mark on a part using an ordinary marker or paint. The application of the mark is done manually, as is the tracking of the part through the plant. The visible mark aids the identification of the part as it travels along the operations and through the conveyor systems. As with the automatic barcoding, the mark is automatically removed during the initial paint operation. Because of the high labor intensity of this method, it is only economical for small studies. It is, however, quite robust to exceptions.

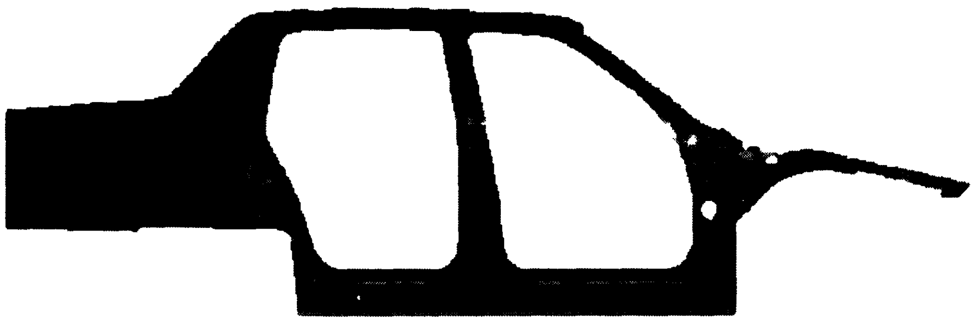


Figure 6 - A mark is painted on a sideframe to manually track it

Benchmarking

Practically every automotive plant uses manual tracking at various times to aid in problem solving. A greater number of plants are beginning to establish automatic methods in their body shops. While it would be impossible to know exactly how many are using automatic tracking, most are starting with the underbodies. One automotive consultant estimated that around 40% of the body shops in North America can track between the underbody and body in white (BIW) (Gretz, 1995). Only a few plants are likely to have the capability to track to any other major sub-assembly.

Many suppliers are offering equipment and solutions to help track sub-assemblies and parts in the body shop. Much of the equipment is designed to interact with the controllers that are already installed in the plants.

How was this Tracking Methodology Implemented?

The goal of this project was to track upstream measurements to the downstream measurements. At the GM assembly plant, this meant tracking as indicated in Figure 7.

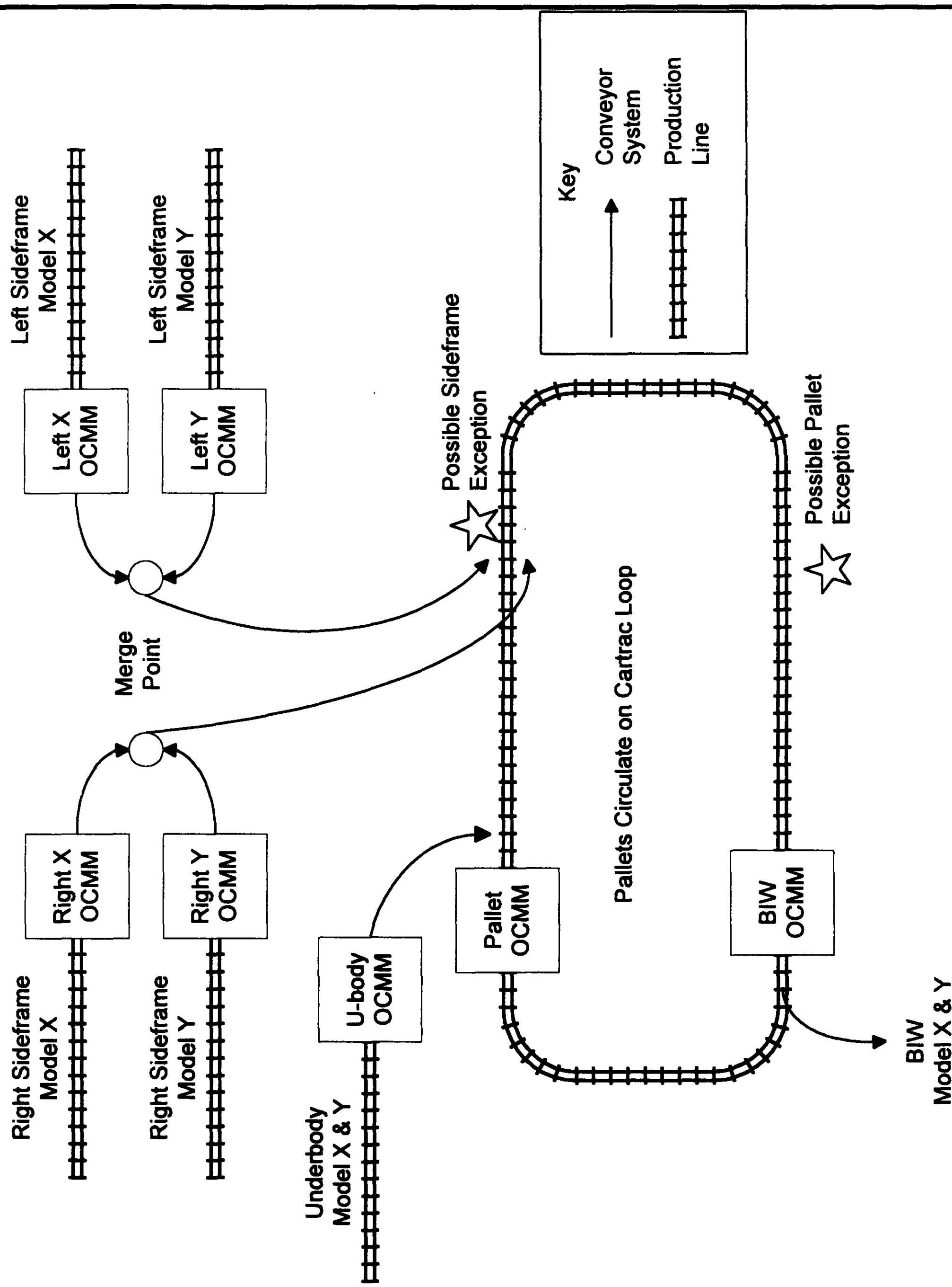


Figure 7 - The goal was to track inputs to outputs

An example Optical Coordinate Measuring Machine (OCMM) record is shown in Figure 8. The first variable of the record is a unique number that identifies the set of measurements. In controllers, such as the BIW and underbody gauges, this number is established by downloading the plant job sequence number (JSN). In the other pallet and sideframe gauges the record ID is sequentially generated by the OCMM controller.

Generic OCMM Record

Record	Date	Time	Aux ID	Measurements
1234567	10-13-94	13:21	0	

BIW OCMM Record

JSN	Date	Time	Pallet	Measurements
5047486	10-13-94	14:27	37	

Figure 8 - An example OCMM record stores information key to tracking

The OCMM gauges also have auxiliary variables that can store information downloaded from the plant information systems. These fields were used to store pallet numbers and sideframe carrier numbers. All OCMM data records are date and time stamped.

Tracking the Underbody to BIW

Electronic tracking between the underbody and BIW at the GM assembly plant already existed. The methodology used was the information system tracking. In that system the underbody is assigned a JSN. This JSN stays with the underbody throughout the body shop. The BIW assumes the JSN of its underbody. The information system passes the JSN to the PLC controllers for downloading to the underbody and BIW OCMM controllers. An example trace between the underbody and BIW is shown in Figure 9.

BIW OCMM Record

JSN	Date	Time	Pallet	Measurements
5047486	10-13-94	14:27	37	

=

Underbody OCMM Record

JSN	Date	Time	Measurements
5047486	10-13-94	10:51	

Figure 9 - The JSN allows tracking between the underbody and BIW

Because an information system is used to do the tracking, the system must be updated to reflect exceptions. When an underbody is removed and another replaced, the material handling system is updated with the new JSN. This update is accomplished by body shop operators who manually enter the changes into a terminal on the shop floor.

Tracking the Pallet to the BIW

The capability to track pallets to the BIW also existed at the GM assembly plant. As described in the background chapter, the pallets are tooling on which the BIW is built. At the assembly plant, there are around sixty pallets that travel around in a loop. Each pallet is identified with a number and has a bar-code attached to it. Sensors read the bar-code as the pallet enters a station and passes the information to the PLC controller. When the OCMM stations measure the pallet and again when they measure the BIW, the stations retrieve the pallet number from the PLC controller. This information is time stamped and stored with the measurements.

The pallet number, date, and time provide a mechanism to track the pallet measurements to the BIW measurements when the pallets circulate on the loop. Because pallets are pulled off the loop for adjustment after both measurements have been taken, there is typically not a problem with exceptions. There are, however, places in the loop where it would be possible to remove a pallet from the loop and these exceptions must be checked. In practice, it is very rare that a pallet would be pulled off the line between the measurements. An example trace between the pallet and BIW measurements is shown in Figure 10.

BIW OCMM Record

JSN	Date	Time	Pallet	Measurements
5047486	10-13-94	14:27	37	

	≤	<	=	

Pallet OCMM Record

Record	Date	Time	Pallet	Measurements
0001005	10-13-94	13:21	37	

Figure 10 - The date, time, and pallet number allow tracking between the pallet and BIW

Tracking the Sideframes to the BIW

Establishing the relationship between the sideframes and the BIW is difficult. When the measurements of the sideframes are made it is not known to which BIW they will become attached. The reason for the initial uncertainty is that there are separate production lines for the sideframes based on the car model. At the assembly plant tracking of sideframes to the BIW was only done for small lots upon request using the hand tracking methodology. One solution for automatic tracking would be to keep track of the sideframes until the JSN of the underbody with which they will merge is known. Another solution would be to label the parts and scan the label when it becomes part of a BIW.

The methodology used in this project was an information system approach. This approach utilized information stored in the plant's material handling system. In particular the tracking focused on the merge system, which had many names: merge point, marriage, and Smarteye (for the sensors that read the bar-codes on the carriers). The merge system identifies upcoming underbodies and sequences sideframes for the appropriate models. The carrier numbers and job sequence numbers were built into the material handling system, but were never stored.

A methodology to track sideframes to the BIW was developed and tested using information from the merge system. The new methodology uses an external database to capture the information stored in the merge system. In addition, the new methodology calls for the downloading of the sideframe carrier numbers into the PLC controller and then into the sideframe OCMM controllers. An example trace between a sideframe and BIW is shown in Figure 11.

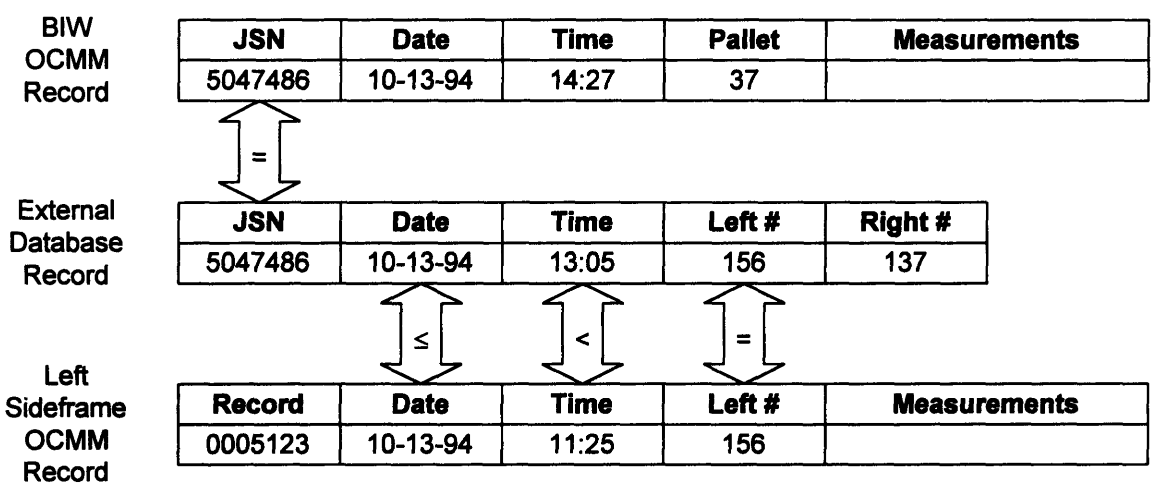


Figure 11 - The JSN, date, time, and carrier numbers allow tracking between the sideframes and BIW

There are exceptions that must be watched when tracking the sideframes to the BIW. Extra sideframes are kept on carts near where they join the underbodies. When a problem sideframe is discovered, it is replaced with a sideframe from a cart. At this time, there is no electronic capture or entry of exceptions as in the BIW. Catching exceptions requires the help of plant floor personnel. When an exception is found, a written record must be kept.

Verification of the Tracking

The verification of the tracking methodologies at the GM assembly plant was performed using only the existing information systems. The external database for the trace was created using data manually taken from the merge system. At the time of the project, the underbody station was not fully installed. Therefore, tracking was only performed between the sideframes, pallet, and BIW stations. The verification trace was run for four days across all shifts. Because the tracking methodology used an information system approach, manual tracking was also performed several times per day to corroborate the results.

In total, 1873 automobile bodies were tracked. The team coordinators were essential in accomplishing a successful trace. Nine exceptions for the week were caught and recorded by the team coordinators and plant personnel. Seven exceptions were due to damaged sideframes while the others were attributable to damaged underbodies.

Implementation

The GM assembly plant is going ahead with plans to implement tracking between the sideframes and BIW. A system architecture was designed to accomplish the tracking and to establish a distributed gauge network on a LAN system. The plan includes some equipment to allow the sideframe OCMM controllers to download the carrier number from the PLC controller. A summary of the purchase request is found in Appendix A and implementation of the plan is underway.

Conclusions

Because of a shortage of statistically trained analysts and a lack of tracking, the data are often only examined on a station by station basis, and then often via an assortment of linear univariate methods. There is a critical need, and much opportunity, to discover important relationships via linear and non-linear multivariate analysis of data across many stations. In this way, better diagnosis of problem processes is enabled, and the prediction of product quality based on process characteristics can be used to better inform continuous improvement and control activities.

For a small cost, this GM assembly plant has tapped into an ability that few plants can emulate: tracking of all major sub-assemblies to the BIW. They are also turning their individual gauges into a true distributed gauging system. This project helped to pave the way to tracking the sideframes to the BIW and culminated in a one-week verification trace.

The solution and methodology for tracking parts are very plant specific. Other plants, even within General Motors, would have difficulties directly adopting the methodologies used here. This is particularly true because tracking was largely accomplished via an information system approach. Critical to the success as well was the involvement of plant personnel. The team coordinators in the body shop were essential in helping to catch exceptions to the tracking procedure.

Chapter 4 - Outliers

What are Outliers?

When measurements are not representative of the "true" value, the measurements can be considered outliers. For example, it would not be representative when the arrow misses the target as in Figure 12.

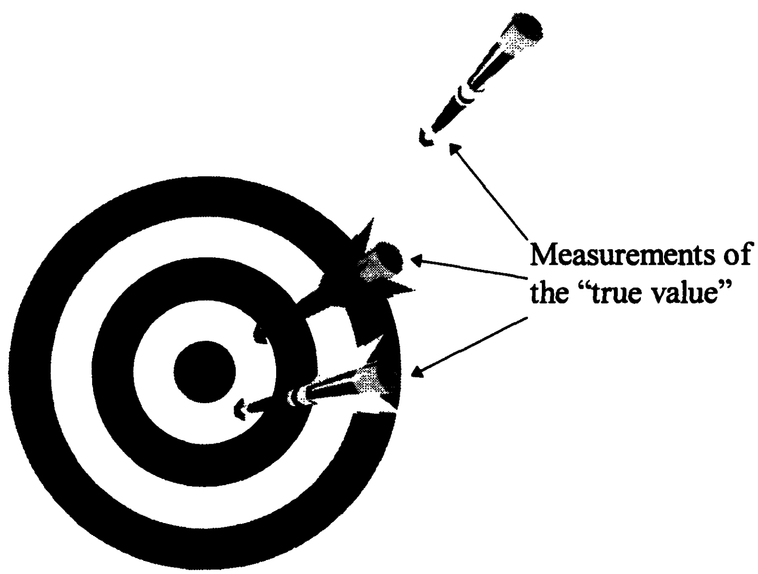


Figure 12 - An outlier occurs when an arrow completely misses the static target

There are biases (sights are off to one side) and errors (stance might be awkward) that affect all of the measurements. Outliers in this example refers to the cases when every once and awhile there is a stray arrow.

If the arrow is way off, it is likely that it will be labeled and called an outlier. A broken arrow could cause an arrow to completely miss the target just as a part that comes unclamped could throw off a measurement. It is possible that the cause of the broken arrows can be identified and that this source of outliers can be eliminated. But as hard as one tries, it is impossible to completely insulate a measurement system from outliers once in a while.

There are many possible causes for outliers in the automotive shop. They include:

- Unwanted measurements - part not in the station
- Dirty parts - grease or paint covering a feature
- Dirty cameras - lens maintenance needs to be performed
- Sub-optimal camera setup - robust setup not performed

This list is not exhaustive, but only represents a sample of what can contribute to measurement error.

Why are Outliers Important?

It is impossible to prevent all outliers and, therefore, some of the data will not be representative of the true values. Unfortunately, not all analysis methods are robust to outliers. Letting the outliers remain in the data in these analyses can lead to the wrong conclusions.

Figure 13 is a graph of 6-sigma, or six times the standard deviation, variations taken by an Optical Coordinate Measuring Machine (OCMM) controller. It is an important graph because it proves the existence of outliers and proves that ignoring that fact can lead to the wrong conclusions. There are two series in the graph: the solid bars represent all the data while the diagonal bars represent data in which the outliers were removed. The continuous improvement indicator (CII) values show that there is a big discrepancy between the two series. The assembly plant uses the CII, which is the 95% ranked 6-sigma point, to summarize the variation for the part. In addition, the complete series indicates a problem at the first two measurements instead of where the highest variation actually is.

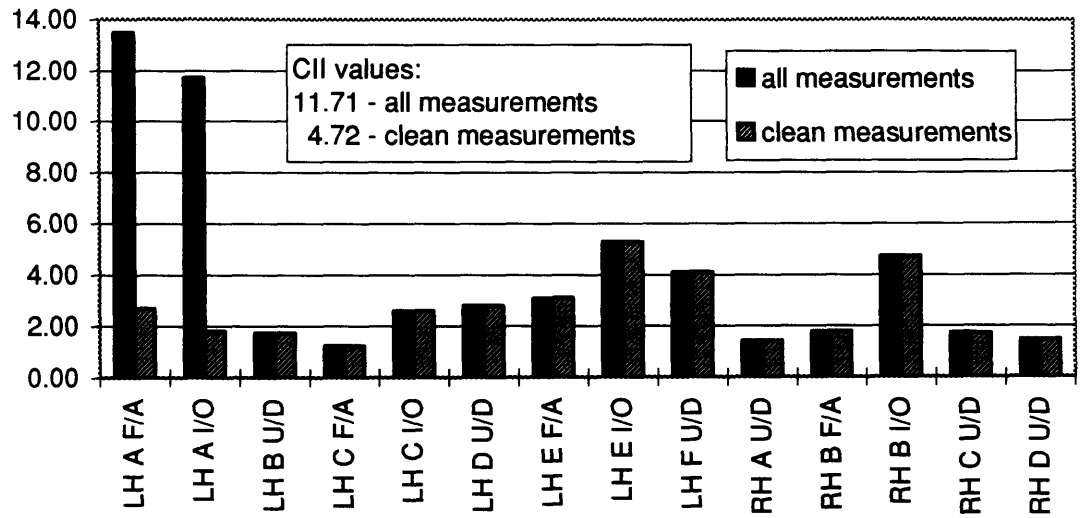


Figure 13 - Outliers in the data can lead to wrong conclusions

How are Outliers Currently Handled?

Outliers are often handled in one of four ways: attacking the root cause, implementing univariate methods, implementing multivariate methods, or leaving the outliers in the data.

Root Cause Identification

Individuals and more often teams are utilized to reduce outliers through root cause identification. When the source can be identified, the plants usually try to eliminate the source of the outliers. There are also times when the teams can uniquely identify a problem but cannot find a root cause. While the search for the cause continues, the identified outliers are systematically screened from the data set.

Variation reduction teams play their biggest role in outlier reduction shortly after installation of the measurement system. Many incorporate gauge repeatability and reliability studies to verify system results. They continue working to improve the system until the study produces acceptable results. A study cannot completely replicate all that will happen in the body shop, however, and the team must constantly be aware that outliers will appear in the data.

Univariate Methods

Univariate methods for outlier detection are based solely on the output values of the variable that are being examined. Boxplots, as in Figure 14, have proved to be quite a good exploratory tool for outliers, especially when several boxplots are placed side by side for comparison (Tukey, 1990). The most striking visual feature is the box, which shows the limits of the middle half of the data. The top of the box is the upper bound value (UBV), below which 75% of the data lies. The line inside the box represents the median. The bottom of the box is the lower bound value (LBV), under which 25% of the data falls. The UBV and LBV determine the height of the box (H), which is called the inner-quartile range. Outlier points are also highlighted when their distance from the box exceeds one and a half times the inner-quartile range. The non-outlier range, often called the whiskers of the box plots, extends to the farthest points of the non-outlier data. The box and whiskers of the boxplots not only show the location and spread of the data, but can be used to indicate skewness as well (McGill *et al.*, 1978).

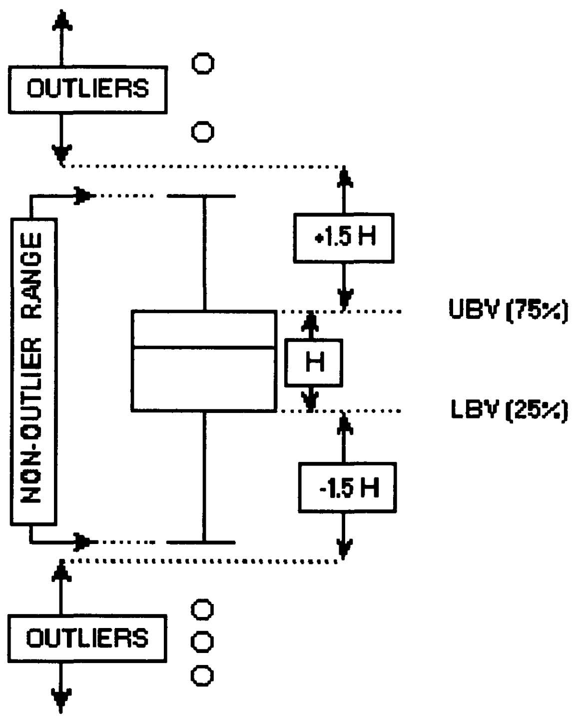


Figure 14 - Boxplots are useful tools for exploring univariate outliers

The boxplots can be used to identify points that may be considered outliers (Velleman and Hoaglin, 1981). The determination of the limits for outliers is based on the middle half of the data points. This range is not likely to contain outliers and is therefore a more robust way to determine the non-outlier range of the data. This method indicates which points are statistically suspected of being outliers, but cannot make the determination as to which points are truly outliers.

Multivariate Methods

Univariate methods are limited in that they only examine the data of one variable independent of the other variables' data. When variables are dependent, univariate methods will have difficulties identifying some outliers. For example, consider the upper and lower control limits for univariate outlier methods in Figure 15. The data point considered will fall within the non-outlier control limits of both variables when using univariate methods. Multivariate methods capture the relationships between the two points and establish a zone that will be able to identify the suspected point. Even when variables are not related, it may be beneficial to use multivariate methods to estimate non-outlier zones that visually would appear as "circles" versus the "square" zones established by univariate methods.

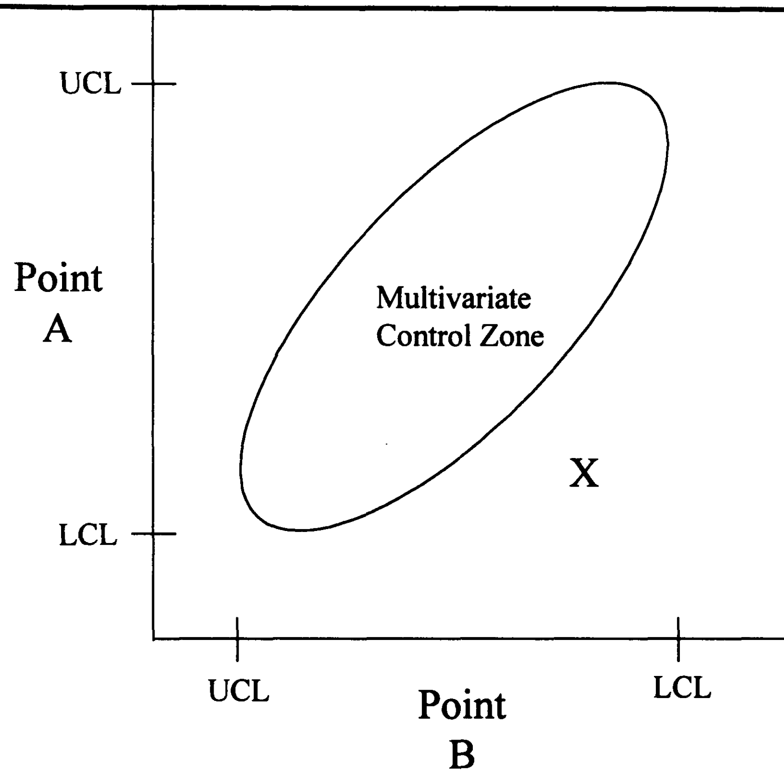


Figure 15 - Multivariate methods increase ability to identify suspect outliers with dependent data

Uncleansed Data

There is a big difference between the methods that statisticians and engineers use and what is practiced on the shop floor. For the most part, industrial practice is to leave the outliers in the data to contaminate the data set. Because it is unreasonable to expect that a gauge as complex as an OCMM gauge to be accurate 100% of the time, this could lead to some erroneous conclusions. All is not lost provided decisions are not made based on one set of measurements. Aggregate studies can provide some protection from outliers for those who want to work with the data without spending the effort to cleanse it from outliers. Aggregate studies provide both means and variances of the data. Comparing these values to historical data will provide some level of confidence in the data.

How were Outlier Methodologies Implemented?

The methodologies used at the GM assembly plant and in this research were the root cause identification and univariate methods. Following are explanations on their implementation and results.

Variation Reduction Teams

The GM assembly plant is fortunate to have active variation reduction teams in the body shop. These teams took responsibility for the OCMM systems and were instrumental in their installation. Each of four teams is headed by the same coordinator, James Clinton. The teams include members from the assembly plant, from GM design, from the General Motors Technical Center, and from the University of Michigan. The teams handle problem solving in four areas of the body shop:

- Sideframe Areas
- Underbody Area

- Cartrac Area (area where the BIW is completed)
- Skidloop Area (area that follows Cartrac)

These teams are largely successful because of the strong leadership and direction of the coordinator. Recognizing that the OCMM systems were only as good as the data they provided, the teams worked hard on the installations using static and dynamic repeatability tests for verification of the data. Their meetings have helped to bring the OCMM stations on-line and have helped to provide good data for variation reduction.

There are many success stories at the assembly plant because of the efforts of the variation reduction team, but they do not stop there. The teams share their ideas on fixing problems of outliers and their ideas on variation reduction with other plants through periodic symposiums.

Univariate Screening

Because it is impossible to prevent all outliers, this research also utilized some univariate methods. Before the methods could be applied to the data, however, some preparatory work was performed. Figure 16 shows the steps performed on the OCMM data to remove outliers.

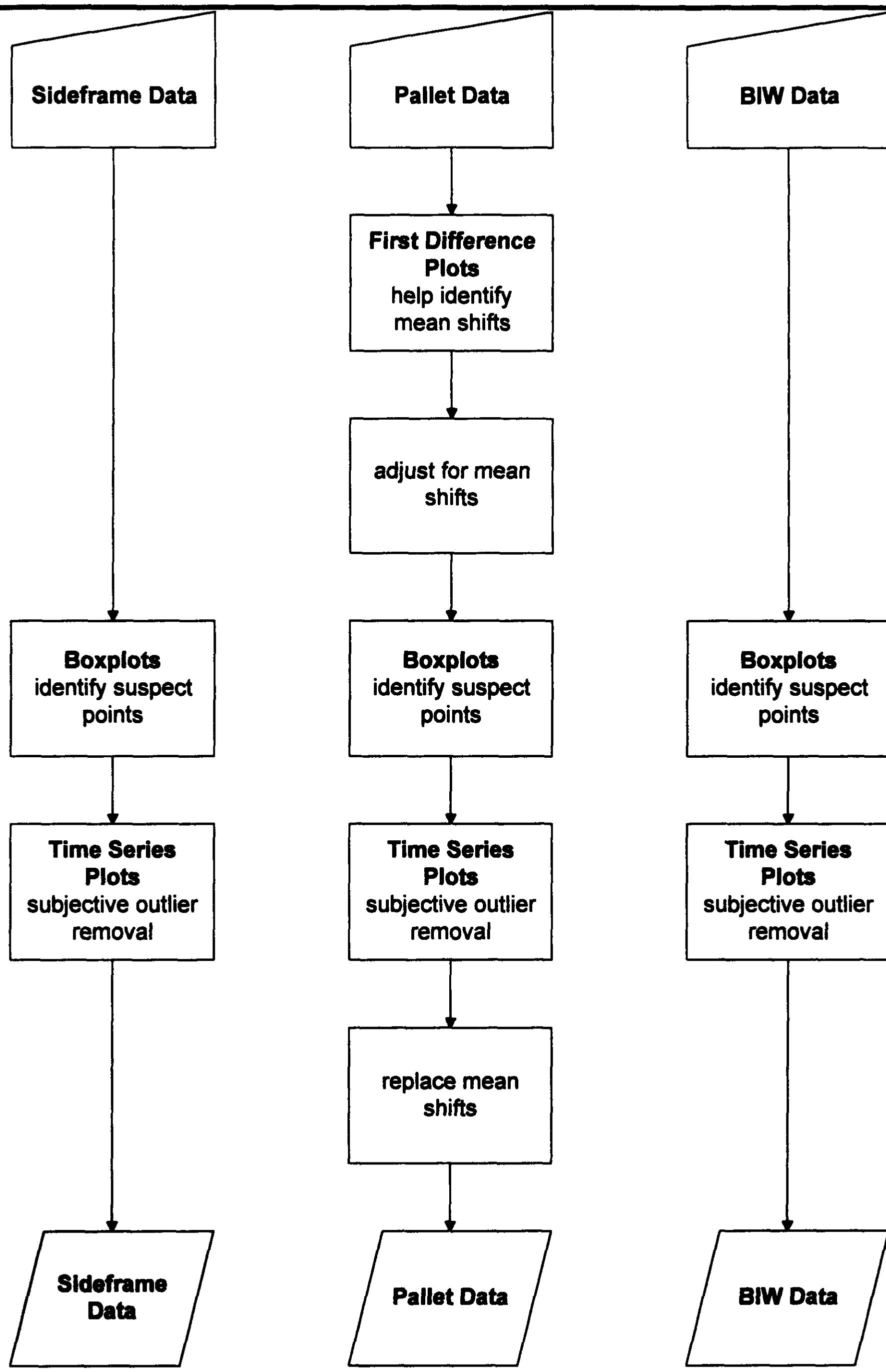


Figure 16 - Univariate methods were used to screen outliers

The data comes from a non-stationary process; adjustments are being made to the process each day. For example, to maintain the pallets within specifications, the toolmakers adjust the locating pins and pads on the pallet with shims. Because many adjustments were being made to the pallets during the data collection period, this data was examined for mean shifts using first difference plots and by looking at the time series plots. The first difference plots, plots of the measurement values less the preceding measurement, specifically were used to find mean shifts of magnitude around 0.25 mm (a typical shim thickness) or greater. These mean shifts were subtracted from the data prior to screening for outliers. Had the mean shifts been left in, the detection of outliers would have been impaired by the larger variation in the data.

Boxplots were implemented on the measurement points to find suspicious data points. Since the data examined did not appear to be skewed, no additional corrective action was needed. The outputs of these analyses were passed onto routines that plotted the variables with respect to time. Points that had been statistically determined as outliers were automatically circled. The ultimate decision to remove a point was subjective. A set of boxplots, Figure 17, and a time series plot, Figure 18, were visual tools that were used to help make that decision.

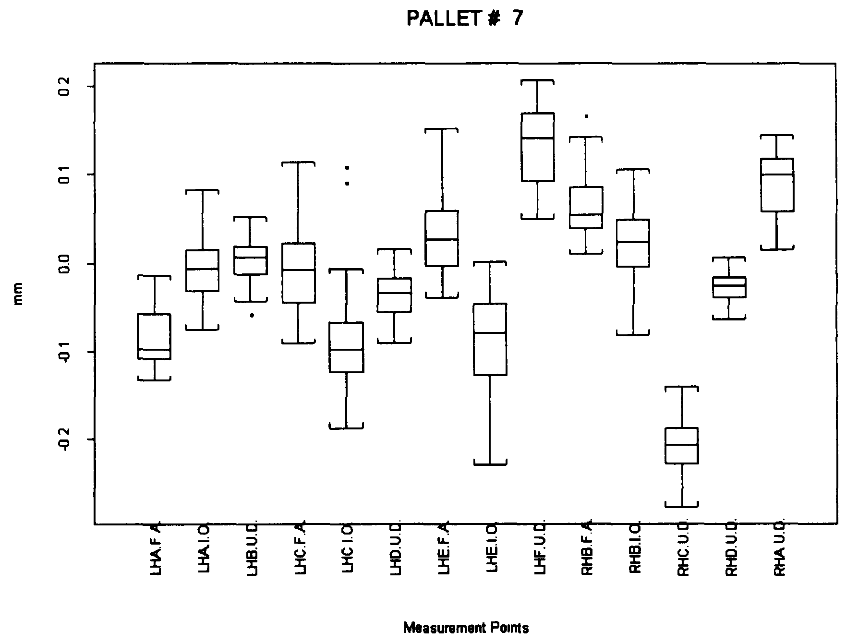


Figure 17 - Boxplots were used to help identify suspect outliers

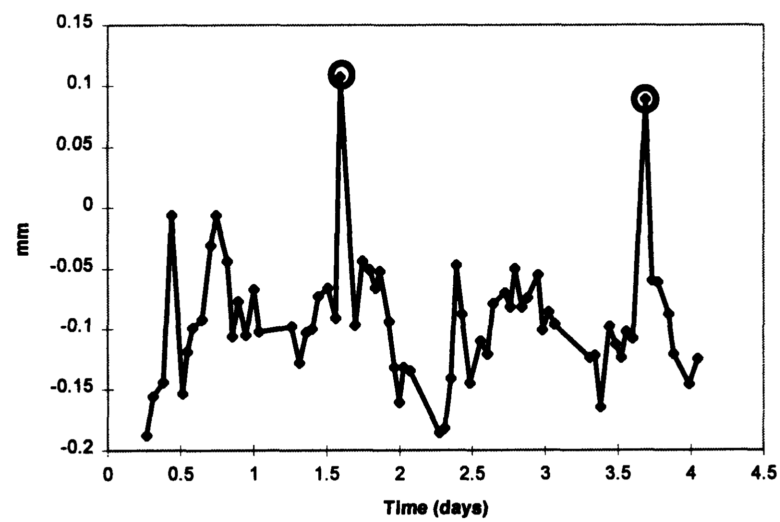


Figure 18 - Time series plot helped in the subjective removal of outliers

Results

Participation on the variation reduction teams and in the symposia was performed as part of the research. One contribution was the identification of the problem shown in Figure 12 and the finding of the root cause. The problem in question arose when a BIW re-circulated on the loop. The pallet system was attempting to measure the pallet with a BIW sitting on it. While the majority of the time the sensors would fail, one sensor would occasionally calculate a measurement based off the OCMM image of the BIW. The variation reduction team quickly discussed and implemented a corrective action to eliminate the problem.

Of the 1873 automobile bodies and parts that were traced in the study, only 532 traces survived the cleansing stages. The other records were thrown out because of either missing data or outliers. Table 3 indicates how many records contained all of the measurements.

Table 3 - Missing data reduced the total number of records

	Pallet	Left Sideframe	Right Sideframe	BIW	Total
Collected	1873	1873	1873	1873	1873
Whole	1866	1791	1848	1284	1227

The cleansing of outliers was as harsh on the data set as was the cleansing of missing data. Table 4 summarizes the results of the univariate screening and subjective removal. The removal of outliers reduced the data set from 1227 records to only 532.

Table 4 - Over one percent of the data points were outliers

	Pallet	Left Sideframe	Right Sideframe	BIW
Data Points Screened	50768	16140	15910	55552
Outliers Found	702	155	258	693
Percentage	1.4%	1.0%	1.6%	1.2%

Conclusions

When an analysis indicates something is wrong, it is probably true. It may be, though, that the problem is with the measurements! It is unreasonable to expect with a complex measurement system with so many variables to have data that is good 100% of the time.

Root cause identification and solutions remain one of the best ways to reduce outliers. A strong and active variation reduction team was instrumental in making root cause identification a success. Because it is impossible to prevent all outliers, some analysis will rely on statistical routines and subjective removal. Univariate methods, while easy to use, have distinct disadvantages over multivariate methods particularly when the variables are dependent.

The test data set was screened for missing data and for outliers using univariate methods. The result was a large reduction in the data set size. Over 34 percent of the traced data records were eliminated because of missing data. While just over one percent of the measurements were outliers, the result was the elimination of another 57 percent of the records.

This cleansing of the data was required because the next analysis step planned was not robust to outliers or missing data. A better treatment to the data set would have been to screen the data set using multivariate methods. In addition, the missing data points and outliers could possibly be replaced using information based on the other data points. The combination of these two would result in the removal of more outliers without the large reduction in data records.

Since the data came from OCMM machines that had recently been installed, the variation reduction teams were still alive and active in root cause analysis activities. It is expected that the percentage of outliers would continue to decrease as more root cause analyses are done and the causes for the "broken arrows" are eliminated one by one.

Chapter 5 - Presentation Error

What is Presentation Error?

When a part comes into a measurement station, it is located before the measurements are taken. The black dots in Figure 19 represent measurements taken on a part. Each time the part is located, however, it is located in a slightly different position. This change in location of the part shows up in the measurements. In other words, every time a new part is measured, the black dots will be in different positions because it is a different part, but also because it is located differently.

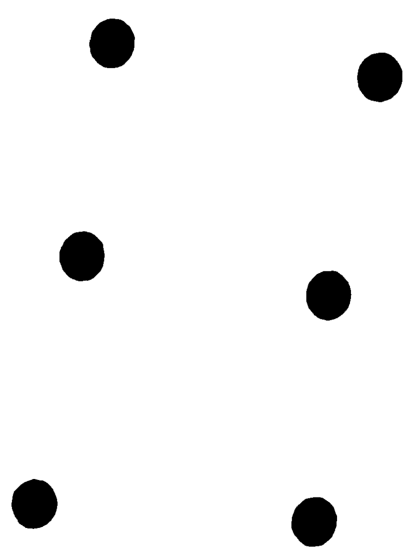


Figure 19 - Each time a part is located, the measurements will be slightly different

The error added to the measurements because of the way the part is located is called presentation error. Perhaps, a better way to see the effect of presentation error is to compare the black dots to the design intent, represented by the open dots in Figure 20.

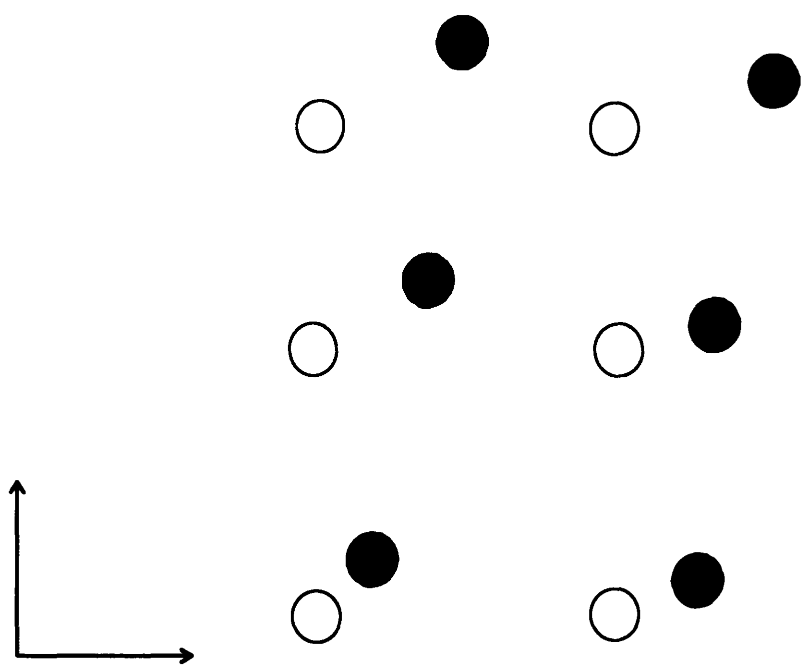


Figure 20 - Comparing measurements to design intent indicates presentation error

Comparing the design intent to the actual measurements indicates that the presentation error is composed of two components: translation errors and rotation errors. If the translation and rotation errors can be determined, they can be removed from the measurements.

Why is Presentation Error Important?

It is important to understand presentation error and how it affects the measurements. Presentation error contributes to the variation contained in the measurements. When the presentation error is a large component of the total variation, it can affect one's ability to draw conclusions from the measurements. Figure 21 shows how the two components of presentation error, translation errors and rotation errors, contribute more to a pallet's variation than the random errors. In this example, the random variation would include the measurement error as well as other unidentified variation.

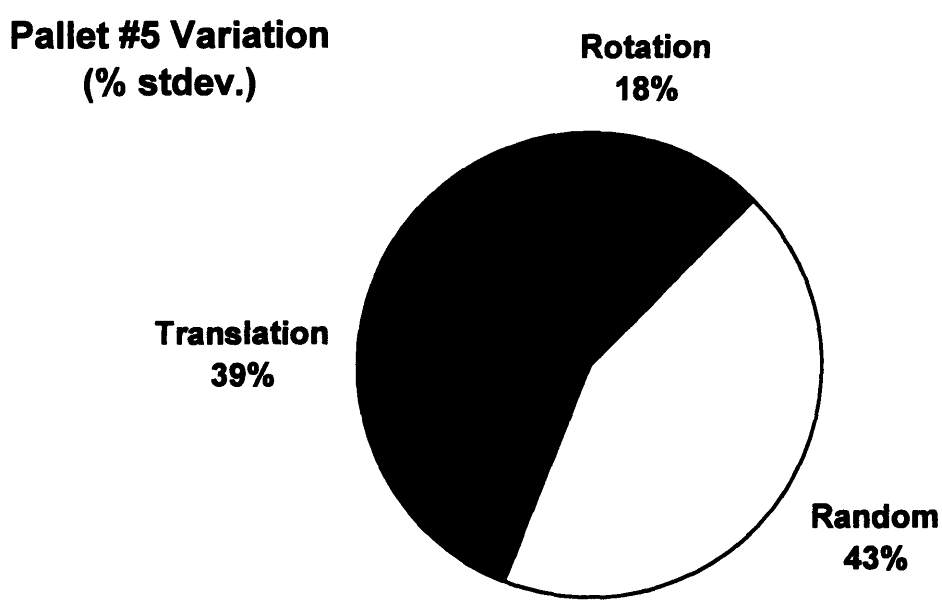


Figure 21 - Presentation error is a significant contributor to the pallet variation

A more vivid example, Figure 22, shows a time series of measurement data from the pallet before and after the presentation error has been removed.

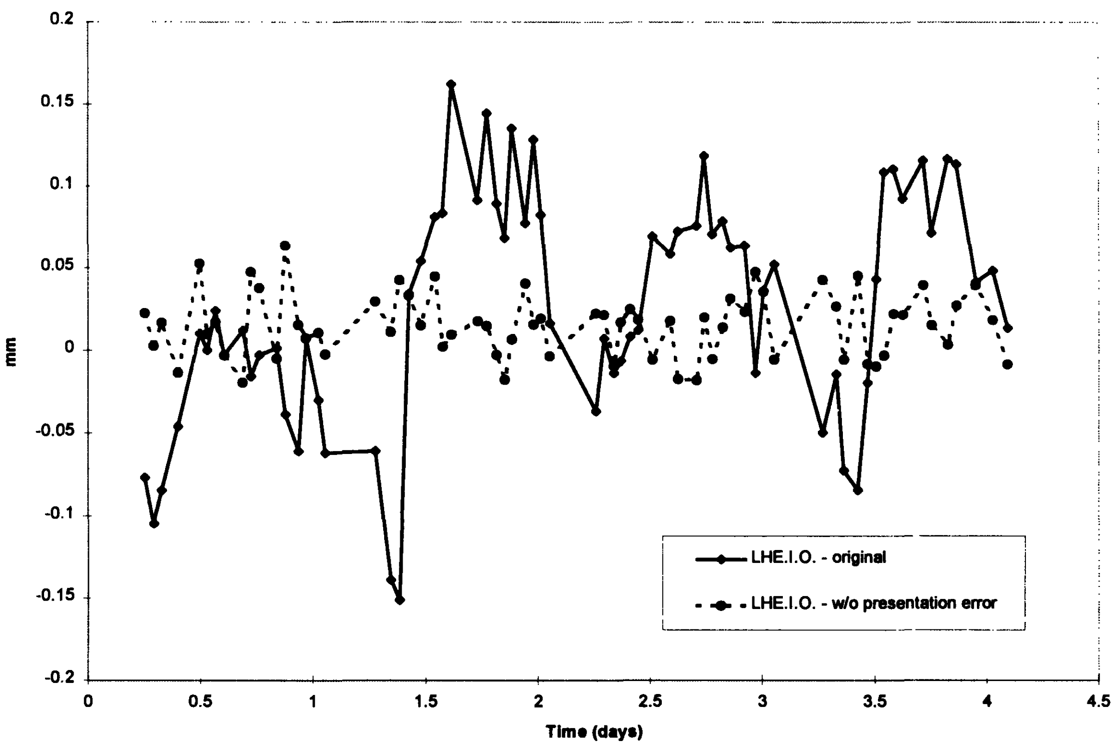


Figure 22 - Removing presentation error reduces the variation

How is Presentation Error Currently Handled?

Installation

For most measurement systems, presentation error is considered only when installing the measurement station. During the installation procedure, the station is calibrated using another measurement instrument. This calibration attempts to account for the biases in the measurements. Following the calibration, a static and dynamic gauge repeatability and reliability study may be conducted to quantify the variation and bias of the measurement system and the variation due to presentation error. Between calibrations, if indeed another calibration is ever conducted, presentation error due to the "play" in the locating fixture is typically not accounted for.

Real Time Correction

A few measurement systems do attempt to adjust for some of this presentation error. For example, a Coordinate Measuring Machine (CMM) measures enough points on the part to establish a 3-2-1 reference frame. The 3-2-1 reference frame becomes the coordinate system to which all the measurements are reported. The Optical Coordinate Measuring Machine (OCMM), as well, can establish a reference frame in a way that they call "visual fixturing" (Perceptron, October 1993). When these two methods are applied correctly, they can reduce the presentation error. These methods do, however, have a few limitations:

- They do not identify the translations and rotations that were removed
- When applied in the wrong place, such as a pallet OCMM station, they throw away important data
- When applied to improper points, such as points with high variation, they may increase the presentation error
- They cannot guarantee that they will find the translation and rotations that minimizes the total mean squared errors

Batch Studies

Part of the reason that presentation error often remains in the data is that research has not provided an efficient method for handling it. Mathematically, it is quick and easy to identify the transformation matrices used in 3-2-1 locating or "visual fixturing," but it is time intensive to identify exactly what the transformation matrix represents.

A few methods have been used to quantify and describe the presentation error. Literature refers to these methods as registration or localization algorithms. Localization is the process of determining the rigid-body translations and rotations that must be performed on the set of points measured on a manufactured surface to move those points into closest correspondence with the ideal design surface (Jinkerson *et al.*, 1993).

General Motors Technical Center has performed off-line analyses to extract presentation error and describe the translations and rotation angle of the errors. Their studies were conducted on two dimensional data sets (Meyer, 1994)

Woncheol Choi and Thomas Kurfess have conducted research on data localization algorithms and minimum zone evaluation for automated inspection (Choi and Kurfess, 1994). Their algorithms are based on a least square fit which can fit to nominal values or determine if a set of measured points can be placed inside a given tolerance zone. While their algorithms determine the transformation matrix, they do not describe the transformation in terms of the translations and rotation angles.

Jinkerson and his colleagues have developed methods to determine the translations and rotation angles of the transformation. Their algorithms, which are based on an iterative technique, determine the translations and Euler angles of rigid-body motion that compose presentation error (Jinkerson *et al.*, 1993). One of the authors, Nicholas Patrikalakis, provided help and advice with the development of the algorithms used in this research.

How was the New Methodology Implemented?

None of the current research provided a solution to the presentation error problem contained in the data from the GM assembly plant. Those methods would fail because they could not handle the combination of all of the following:

- 3-D translations and rotations
- Limited number of data points
- Some measurement points were 2-D points, while other points were 1-D points

The methodology, developed in conjunction with Vikas Sharma, is a new non-iterative solution to the localization problem. It determines the rigid-body translations and rotations that will bring the measured data points as close to the design nominal values as possible. The methodology determines the three translations and the three Euler angle rotations even when provided limited data, most of which do not contain measurements on all three coordinates.

The methodology can be divided logically into four sections: data preparation, transformation matrix estimation, rotation matrix interpretation, and model selection.

Data Preparation

The data obtained from the assembly plant was processed prior to running it through the data localization algorithm. Figure 23 represents the processes that were performed on the data. The first two processes removed records with missing data and records with outliers. These were performed as discussed in Chapter 4. The remaining processes converted the data into a global Cartesian coordinate system.

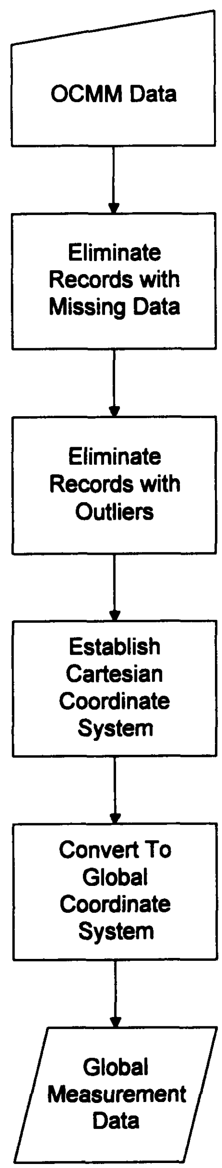


Figure 23 - The data was pre-processed prior to the data localization algorithm

The records were eliminated because the algorithm relies on least squares linear regression techniques. These techniques are not robust to outliers, which can have a big impact on the linear coefficient results. There is room for improvement on robustness to outliers when there are plenty of data points. Alternative techniques for regression based on subsets of the data could replace the linear regressions in future applications of the algorithm. Robustness to missing data has already been improved through the efforts of Vikas Sharma, but these improvements were not tested on the data from the GM assembly plant (Sharma, York, *et al.*, 1995).

The data obtained from the OCMM stations were retrieved in deviation from nominals based on the automotive coordinate system. The automotive coordinate system consists of four axes: fore/aft, up/down, in/out left side, and in/out right side. The in/out axes are co-linear, but have opposite sign conventions. Conversion to a three axis Cartesian coordinate system was accomplished by changing the sign of all in/out measurements on the left side of the body. Conversion to a global coordinate system was achieved by adding the global design nominals to the measurements. Alternatively, these conversions could have been performed by routines imbedded in the OCMM controller.

Transformation Matrix Estimation

The data localization algorithm uses linear regression techniques to estimate values for un-measured components and to estimate the rotation matrix. If the random error in the measurements has a distribution close to normal, a least squares fit will work well (Choi and Kurfess, 1994). The flow diagram of these processes is shown in Figure 24.

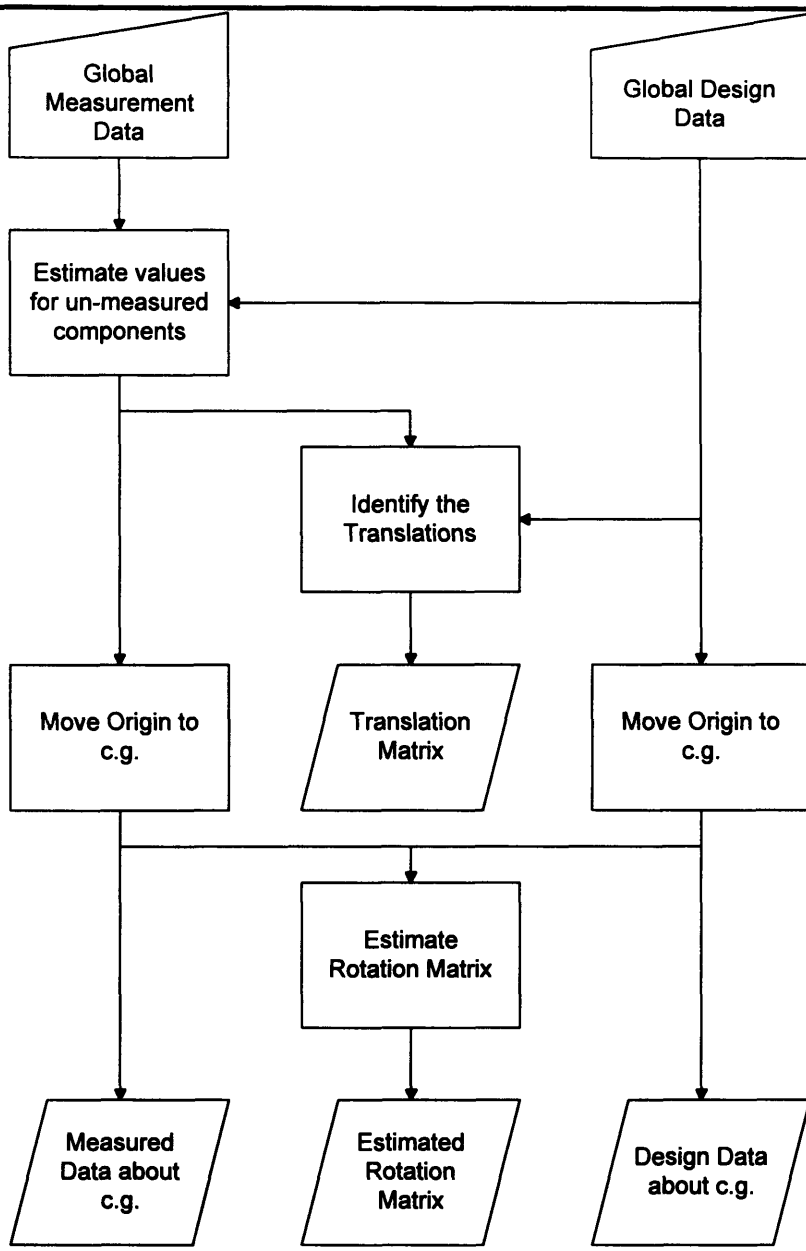


Figure 24 - Linear regression is used to estimate the rotation matrix

The first process in the algorithm is to estimate values for un-measured components. Many of the data points from the OCMM stations only provide measurements in two or one dimensions. The other dimension(s) could be assigned the global design intent, however, a better estimate would also use information stored in the other data measurements.

For simplicity, consider a two dimensional case. A part enters the station and measurements are taken. Some of the measurements are in two dimensions, while the others are only in one dimension. In this case, there are enough data points in each dimension to form Equations 1 and Equations 2. These equations relate the measurements, found on the left hand side of the equation, to the primed design intent values on the right. A least squares fit on these equations will provide the coefficients and constants that allow an estimation of the un-measured values based on the global design intent. For a two dimensional case, at least three measurements would be required on each axis.

$$\begin{gathered} x_1 = a_{11}x_1' + a_{12}y_1' + a_{13} \\ \vdots \\ x_n = a_{11}x_n' + a_{12}y_n' + a_{13} \end{gathered}$$

Equations 1

$$y_1 = a_{21}x_1' + a_{22}y_1' + a_{23}$$
$$\vdots$$
$$y_n = a_{21}x_n' + a_{22}y_n' + a_{23}$$

Equations 2

The next process in the algorithm calculates the translation components. These translations are based on the complete set of data that includes measured and estimated un-measured data. The result is obtained by Equations 3, which subtracts the design intent from the complete set of data.

$$a_{13} = \overline{x} - \overline{x'}$$
$$a_{23} = \overline{y} - \overline{y'}$$

Equations 3

The origins are moved to the centers of gravity for both the complete set and the design data so that the rotation will occur about the same physical point. A linear regression about the centers of gravity provides an estimate of the rotation matrix. The relationship is noted in matrix form by Equation 4.

$$\begin{bmatrix} x - \bar{x} \\ y - \bar{y} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x' - \bar{x}' \\ y' - \bar{y}' \end{bmatrix}$$

Equation 4

Rotation Matrix Interpretation

Interpretation of the rotation matrix in a two dimensional case is easy because the rotation can only occur about the axis perpendicular to the plane. The relationship is given by Equation 5. Based on the linear regression results, the angle of rotation can be determined from the four estimates of gamma. The calculated Euler angle, gamma, is used to generate the actual rotation matrix, which is constructed as indicated on the right side of the equation.

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} \cos\gamma & \sin\gamma \\ -\sin\gamma & \cos\gamma \end{bmatrix}$$

Equation 5

In three dimensions, there are three axes about which to rotate. Rotations about each of these axes are given by Equation 6, Equation 7, and Equation 8.

$$R_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\gamma) & \sin(\gamma) \\ 0 & -\sin(\gamma) & \cos(\gamma) \end{bmatrix}$$

Equation 6

$$R_y = \begin{bmatrix} \cos(\beta) & 0 & -\sin(\beta) \\ 0 & 1 & 0 \\ \sin(\beta) & 0 & \cos(\beta) \end{bmatrix}$$

Equation 7

$$R_z = \begin{bmatrix} \cos(\alpha) & \sin(\alpha) & 0 \\ -\sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Equation 8

The complication is that the three rotations can be combined into an equivalent rotation matrix, an example of which is Equation 9. The order of rotations matters, however, as shown in Equation 10.

$$R_{xyz} = R_x R_y R_z$$

Equation 9

$$R_x R_y R_z \neq R_x R_z R_y \neq R_y R_x R_z \neq R_y R_z R_x \neq R_z R_x R_y \neq R_z R_y R_x$$

Equation 10

There is an infinite number of combinations of rotations when axes are combined and repeated. The trick is to come up with the proper order and proper angles to replicate the rotation matrix generated by the least squares fit. It is only possible because there are exactly twenty-one combinations that could represent the shortest path for removing the presentation error. These combinations are shown in Table 5. They are derived by taking three axes at a time without repeats, two axes at a time while repeating the first axis, two axes at a time without repeats, and one axis at a time.

Table 5 - There are twenty-one "shortest path" combinations

Rxyz	Rxzy	Ryxz	Ryzx	Rzxy	Rzyx
Rxyx	Rxzx	Ryxy	Ryzy	Rzxz	Rzyz
Rxy	Rxz	Ryx	Ryz	Rzx	Rzy
Rx	Ry	Rz			

Each of these twenty-one models is fitted to the estimated rotation matrix. The process, shown in Figure 25, results in the generation of the Euler angles, rotation matrix, measurement residuals, and error values for each model.

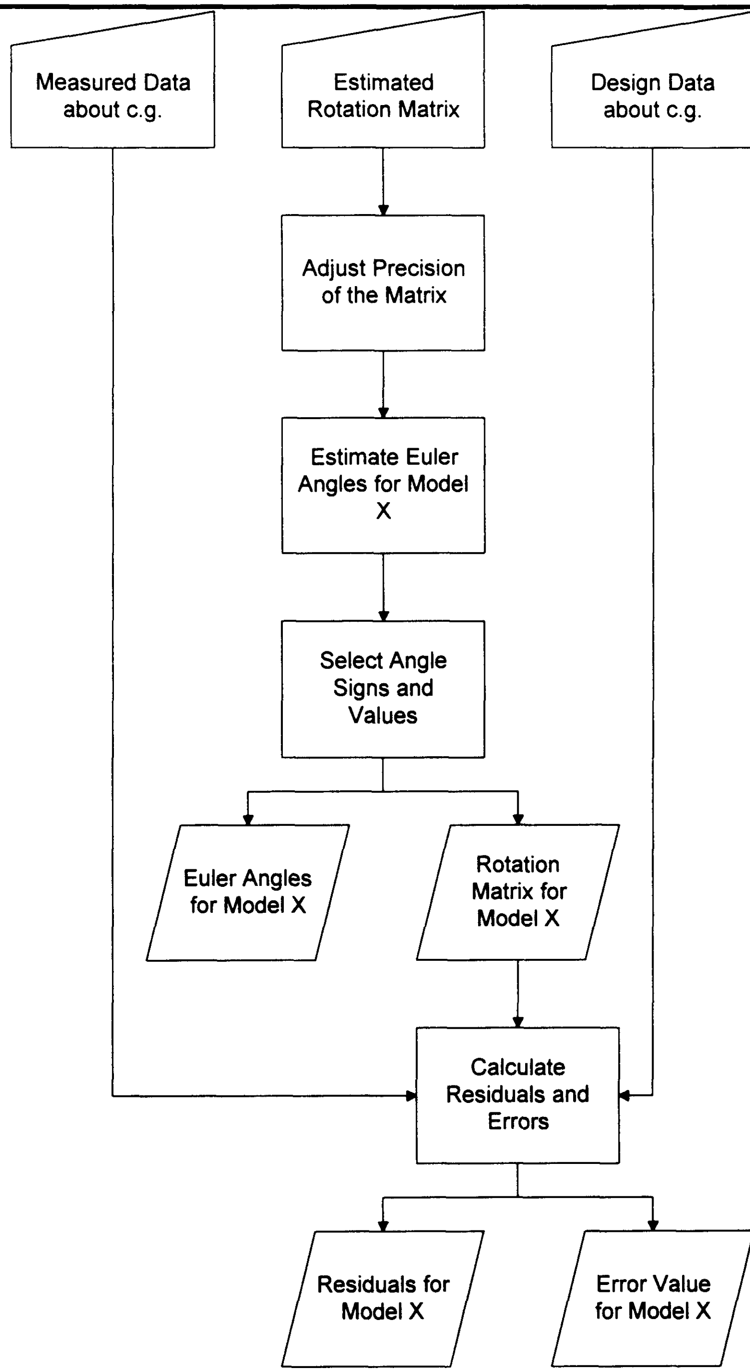


Figure 25 - Each of the twenty-one models are fitted to the estimated rotation matrix

The first process when applying a model is to adjust the numbers in the estimated matrix. Values that are less than 1 e^{-7} are set to zero. Values that are within 1 e^{-7} of one are set to one. This modification, while not affecting the result, helps prevent the routines from making insignificant calculations. The OCMMs report measurements to the 0.001mm (regardless of significance). Because the OCMMs measure parts that span nearly 4000mm, the angles that could be recorded in the data are of a magnitude greater than 1 e^{-7}.

The second process is to estimate the Euler angles for the model. Single axis rotations are easily fitted because there is a one to one map between the estimated matrix element to an element containing a single trigonometric function of the angle. Other combinations, such as Equation 11, do not have the one to one mapping.

$$R_{xyz} = \begin{bmatrix} \cos(\alpha)\cos(\beta) & \cos(\beta)\sin(\alpha) & -\sin(\beta) \\ \cos(\alpha)\sin(\beta)\sin(\gamma) - \cos(\gamma)\sin(\alpha) & \cos(\alpha)\cos(\gamma) + \sin(\alpha)\sin(\beta)\sin(\gamma) & \cos(\beta)\sin(\gamma) \\ \cos(\alpha)\cos(\gamma)\sin(\beta) + \sin(\alpha)\sin(\gamma) & \cos(\gamma)\sin(\alpha)\sin(\beta) - \cos(\alpha)\sin(\gamma) & \cos(\beta)\cos(\gamma) \end{bmatrix}$$

Equation 11

The non-iterative approach uses trigonometric relationships between various elements in the estimated matrix and the fitted model. This relationship is demonstrated by Equation 12. The relationships for this model produce estimates for the three angles as shown in Equations 13, Equations 14, and Equations 15. The equations used in each rotation model are not exhaustive, other equations could be written using other combinations of the matrix elements. Enough equations must be developed for each rotation model to provide good estimates of the angles. Testing of the equations against test data is essential to assure that the equations are sufficient for determining the angles.

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = R_{xyz}$$

Equation 12

$$\alpha_1 = \sin^{-1}\left(\frac{a_{31}a_{23} - a_{21}a_{33}}{\sqrt{1-a_{13}^2}}\right)$$

$$\alpha_2 = \cos^{-1}\left(\frac{a_{22}a_{33} - a_{32}a_{23}}{\sqrt{1-a_{13}^2}}\right)$$

$$\alpha_3 = \sin^{-1}\left(\frac{a_{12}}{\sqrt{1-a_{13}^2}}\right)$$

$$\alpha_4 = \cos^{-1}\left(\frac{a_{11}}{\sqrt{1-a_{13}^2}}\right)$$

$$\alpha_5 = \tan^{-1}\left(\frac{a_{12}}{a_{11}}\right)$$

Equations 13

$$\beta_1 = \sin^{-1}\left(\frac{a_{21}a_{23} + a_{31}a_{33}}{a_{11}}\right)$$
$$\beta_2 = \sin^{-1}\left(\frac{a_{22}a_{23} + a_{32}a_{33}}{a_{12}}\right)$$
$$\beta_3 = \sin^{-1}\left(\frac{a_{11}a_{21} + a_{12}a_{22}}{a_{23}}\right)$$
$$\beta_4 = \sin^{-1}\left(\frac{a_{11}a_{31} + a_{12}a_{32}}{a_{33}}\right)$$
$$\beta_5 = \sin^{-1}(-a_{13})$$

Equations 14

$$\gamma_1 = \cos^{-1}\left(\frac{a_{11}a_{22} - a_{21}a_{12}}{\sqrt{1 - a_{13}^2}}\right)$$
$$\gamma_2 = \sin^{-1}\left(\frac{a_{31}a_{12} - a_{11}a_{32}}{\sqrt{1 - a_{13}^2}}\right)$$
$$\gamma_3 = \sin^{-1}\left(\frac{a_{23}}{\sqrt{1 - a_{13}^2}}\right)$$
$$\gamma_4 = \cos^{-1}\left(\frac{a_{33}}{\sqrt{1 - a_{13}^2}}\right)$$
$$\gamma_5 = \tan^{-1}\left(\frac{a_{23}}{a_{33}}\right)$$

Equations 15

The next process determines the sign and values for the three angles. Because inverse trigonometric functions are used in the estimation of the angles, the values must be corrected. Depending on the quadrant, the trigonometric functions will yield different values, as shown in Table 6.

Table 6 - An example from each quadrant shows the sign and value problem

	$\theta = 45°$	$\theta = 135°$	$\theta = -135°$	$\theta = -45°$
$\sin^{-1}(\sin(\theta))$	45°	45°	-45°	-45°
$\cos^{-1}(\cos(\theta))$	45°	135°	135°	45°
$\tan^{-1}(\tan(\theta))$	45°	-45°	45°	-45°

When angle estimates were obtained using at least two different inverse trigonometric functions and do not fall on an axis, the quadrant can be determined from Table 6. The estimates are adjusted according to Table 7. The median estimation for each angle is selected.

Table 7 - Adjustments are made according to the quadrant

	Quadrant I	Quadrant II	Quadrant III	Quadrant IV
$\theta = \sin^{-1}(\sin(\,))$	θ	$\theta + 90°$	$\theta - 90°$	θ
$\theta = \cos^{-1}(\cos(\,))$	θ	θ	$-\theta$	$-\theta$
$\theta = \tan^{-1}(\tan(\,))$	θ	$\theta - 180°$	$\theta + 180°$	θ

The final angles are put back into the model's matrix equation, such as Equation 11. The inverse transformation is applied to the measurements and the residuals calculated. The mean square error for the model is the three dimensional residual deviation from design intent.

Model Selection

After each of the twenty-one models has been fitted to the estimated rotation matrix, the errors and angles are compared. If more than one model reduces the mean square error below a threshold, the model with the least total rotation for the body is chosen. Otherwise, the model with the least mean square error is selected.

Results

This methodology was developed working in conjunction with Vikas Sharma at MIT. The algorithm was successfully tested on generated data with and without noise. In addition, it was compared to an iterative approach using data from a propeller blade and a sinusoidal surface (Sharma, York, *et al.*, 1995).

The results of the algorithm were remarkable for a single pallet, which is re-measured each time it travels around the production loop. Figure 21 visually shows how presentation error accounts for more than half of the pallet variation. The translation error accounted for 39% while the rotation errors accounted for 18%. Separating the presentation error from the measurements increased the signal to noise ratio, as depicted in Figure 22.

As expected, when part to part variation is added the contribution of presentation error is less, but still significant. For the data collected during Chapter 3, 8.9% of the mean square error for the pallets was reduced by removing the presentation error. The body in white mean square error was reduced by 7.1%.

Conclusions

A new methodology for removing and identifying presentation error was developed because no current algorithms could simultaneously address all of the limitations imposed by the data from the GM assembly plant. These limitations included the limited number of data points, the three dimensional translations and rotations, and that few of the data points had measurements in all three dimensions.

The routines were successfully implemented on the data from the OCMM gauges. The results indicated that a significant amount of variation was attributable to the presentation error. Although the methodology was applied to the OCMM data, it is device independent and was verified using generated data and data sets from other sources.

There are improvements that can be made to the methodology. Some progress in the areas of missing data and angle determination has occurred since the algorithm was developed for the GM data (Sharma, York, *et al.*, 1995). When the number of data points is large, such as with the body in white measurements, improvements in robustness to outliers should be possible by using more robust linear regression techniques. Future research could improve the speed of the algorithm by selectively fitting only a few of the twenty-one models based on characteristics of the estimated rotation matrix.

Chapter 6 - BIW Ruler

What is a BIW Ruler?

It is difficult to describe a body in white (BIW) ruler without telling why it is important. Figure 26 shows data taken from an Optical Coordinate Measuring Machine (OCMM) at a GM assembly plant. If one data point in the figure is selected, its position alone is not important.

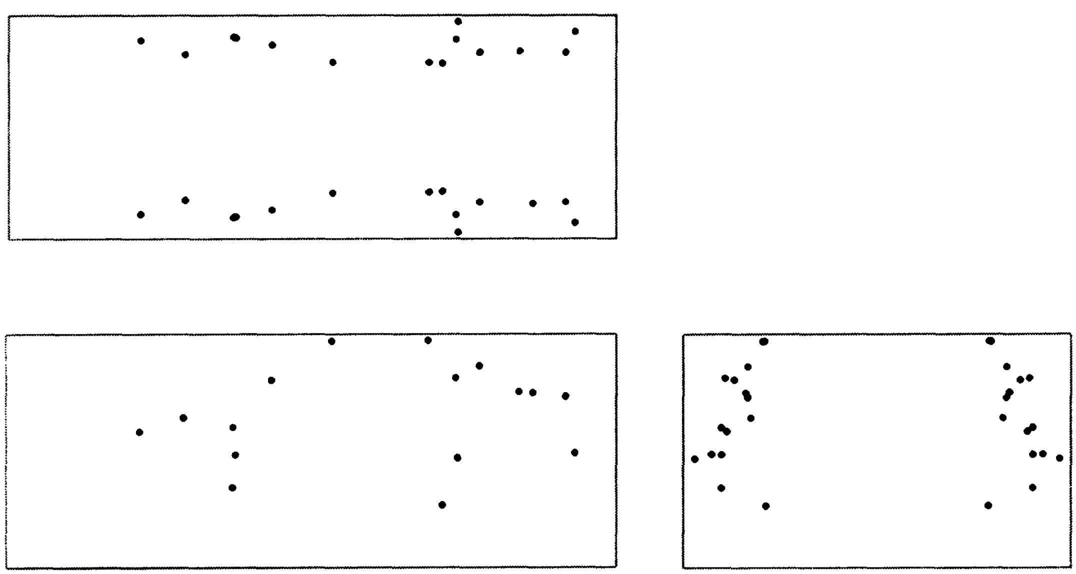


Figure 26 - A data point is one of many others

What is important are the relationships of that point to the other data points. Figure 27 shows that these measurements collectively define the BIW. A BIW ruler is a way to measure the relationships of the data.

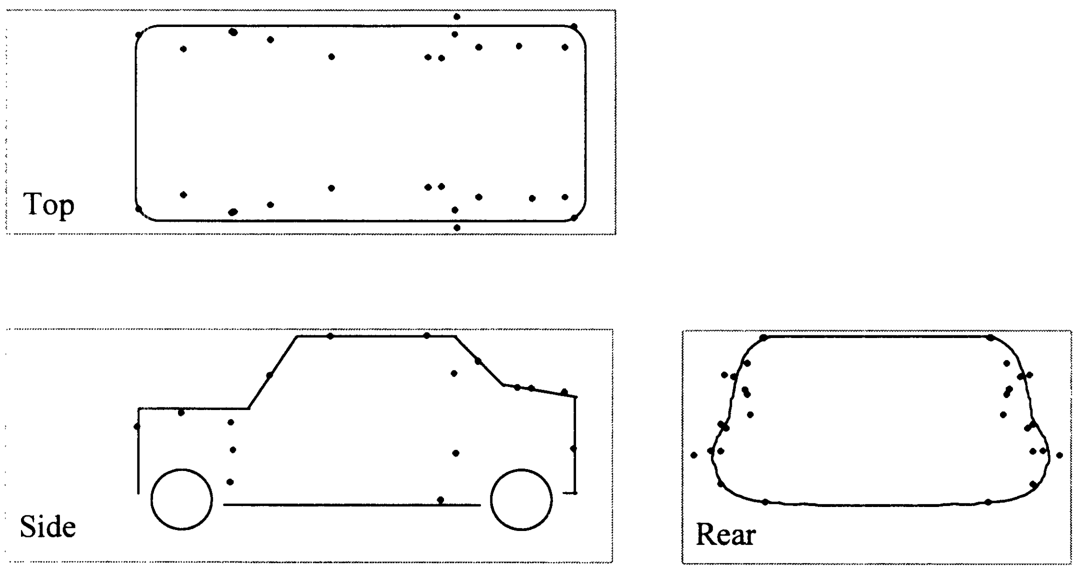


Figure 27 - The data points collectively define the BIW

Why is a BIW Ruler Important?

A BIW ruler is important because it attempts to describe the relationships and behavior of the data. It attempts to answer questions like:

- Is the width of the windshield opening too small?
- Is the width of the roof at the front larger than the width at the back?
- Is the back of the deck opening (trunk opening) off to one side?

Answers to these questions are useful both for proactive control and for data analysis.

The relationships help in the understanding of how the BIW is building. Being able to watch key relationships helps simplify control in the body shop. The end goal is to measure the relationships that can best be used to understand, improve, and control the process in order to produce a high quality automobile for the customer.

Relationships can help in some analyses by reducing the dimensionality of the data set. In the body shop, more data variables are likely measured on the BIW than the sum of all upstream measurement variables. Relationships reduce the complexity of the analysis while retaining much of the information stored in the data. For example, an analysis could be performed to find which upstream variables have an effect on the relationships in the deck opening. The deck opening relationships are important because they indicate how the deck lid (trunk) will eventually fit (or how expensive it will be to make the deck lid fit).

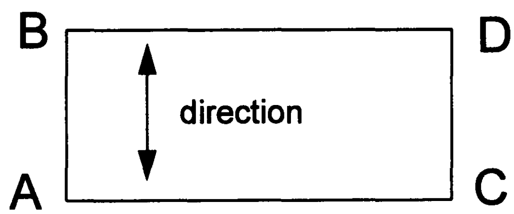
How is a BIW Ruler Currently Handled?

A lot of plants resort to relationships only when reacting to problems. When final assembly comes to the body shop and tells them that the back window is leaking, the body shop will go back and check the cross-car measurements to see if they are building within tolerances. During the rest of the time, however, few relationships are used. When relationships are used, they are typically only difference equations. For example, a difference equation would be the distance between two holes. Most fail to capture any of the more advanced relationships.

Relationships are not new to General Motors. The GM assembly plant has been requesting cross-car relationships for the OCMM controllers since November 4, 1993 (Clinton, December 1994). There is evidence as well that General Motors has used advanced relationships in the past. In February 1981, relationships were generated to look at widths, tapers, and skewness at another GM facility (Harvey, 1982). They utilized relationships to reduce the number of variables that the tooling people had to examine.

How was this Methodology Implemented?

The methodology for capturing the relationships is similar to the methodology implemented at the other GM facility, though it was developed independently. For a zone, such as the roof, four fundamental behaviors are captured: front width, back width, taper, and skewness. These relationships are shown in Figure 28.



Measurements: A,B,C,D

Design: A',B',C',D'

$$y_1 = \frac{(A-B)-(A'-B')}{A'-B'} \qquad y_2 = \frac{(C-D)-(C'-D')}{C'-D'}$$

$$y_3 = \frac{[(A-B)-(A'-B')]-[(C-D)-(C'-D')]}{\frac{1}{2}[(A'-B')+(C'-D')]}$$

$$y_4 = \frac{[(A+B)-(A'+B')]-[(C+D)-(C'+D')]}{[(A'-B')+(C'-D')]}$$

Figure 28 - Four equations capture the fundamental relationships

The first relationship captures the width of the roof near the windshield. The second relationship describes the opposite width of the roof. These two width relationships are equations that compare them to design intent and express the result as a percentage. A third equation relates how the two widths relate to each other. If one is too big and the other too small then a taper results. A taper translates to "A" or "V" gaps when other parts are added to the BIW. A fourth equation describes the behavior when the roof is skewed. The interpretations of these equations are indicated in Figure 29.

relationship	meaning	+	-
1	% width	too wide	too narrow
2	% width	too wide	too narrow
3	% A or V gap		
4	% skew		

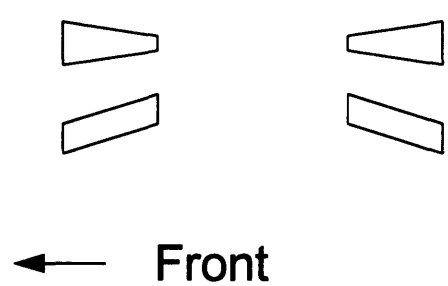


Figure 29 - The relationships have physical interpretations

The relationships are expressed as percentages so that they can be compared to relationships in other areas of the BIW. The relationships can be organized into a hierarchical system as in Figure 30.

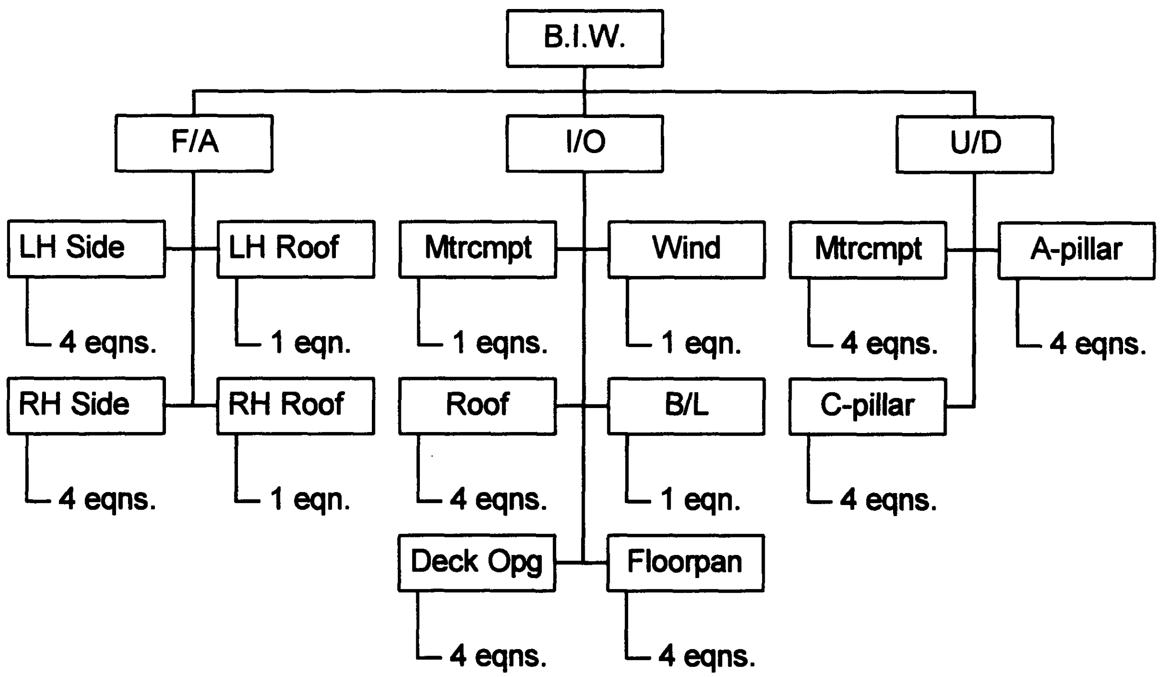


Figure 30 - The relationships are related to other relationships

An examination of Figure 30 reveals that not all zones of the BIW have four equations. In some zones, there are not enough measurements to form four equations. Figure 31 graphically indicates where the relationships exist in the in/out direction. The roof, floorpan, and deck opening each have four equations while the engine compartment, windshield, and back window each have one width equation.

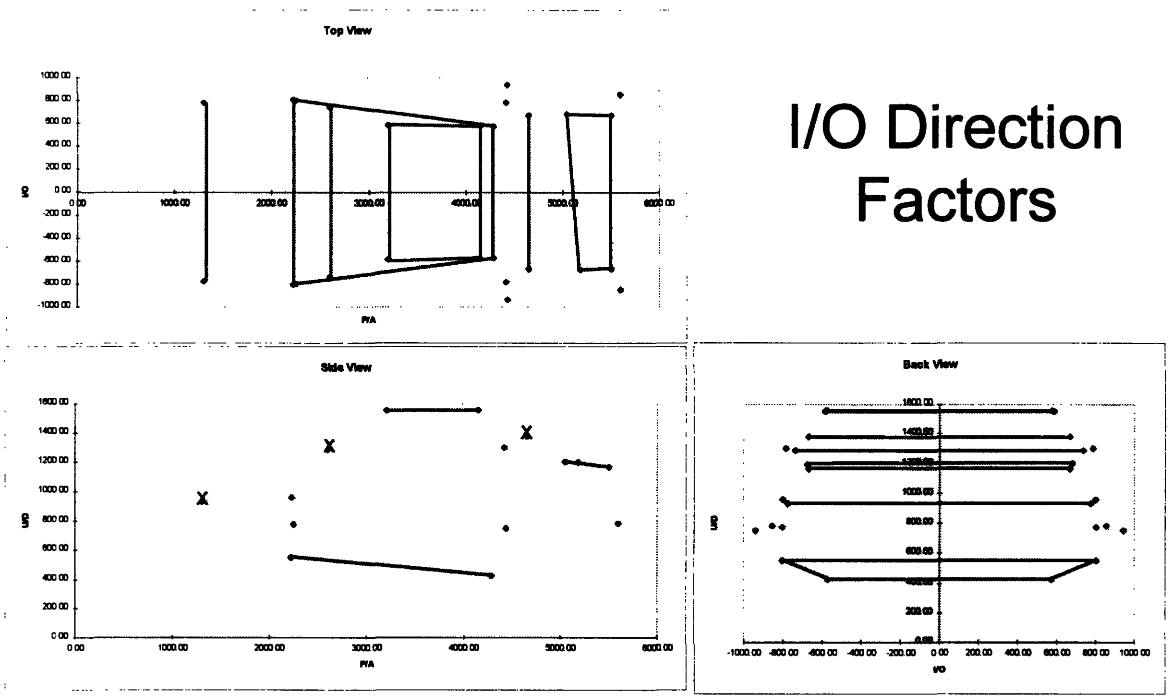


Figure 31 - A few in/out zones have four equations while others have one equation

Implementation of these relationships can be performed off-line or on-line. The GM assembly plant asked for one set of relationships to be programmed into the OCMM controller for evaluation. The roof zone for the in/out direction was chosen and programmed on December 13, 1994. Examples of the reports available along with interpretations are found in Appendix B. A weekly feedback mechanism on the relationships was incorporated into the variation reduction meetings held in the body shop.

Conclusions

Relationships between data points are important because they help in the understanding of how the BIW is building. Being able to watch key relationships helps simplify control and variation reduction in the body shop. Relationships help in comparing different zones on the BIW as well as in comparing different BIWs.

Relationships also help during some analyses by reducing the dimensionality of the data set. This reduction is useful in the body shop because there are typically more variables measured in the downstream BIW gauge than in the upstream gauges. Relationships reduce the complexity of the analysis while retaining much of the information stored in the data.

It is troubling to see that so few relationships are being used in the body shop. There is much to be gained by monitoring significant relationships. More troubling is that each time advanced relationships are used, the “wheel” seems to be re-invented.

Advanced relationships were implemented on the OCMM controller at the GM assembly plant to monitor the roof widths, taper, and skewness. Weekly feedback on the relationships during the initial months proved positive. The plant requested additional relationships to be programmed for the engine compartment and deck opening.

Chapter 7 - Conclusions

Because of a shortage of statistically trained analysts and a lack of tracking, the data are often only examined on a station by station basis, and then often via an assortment of linear univariate methods. There is a critical need, and much opportunity, to discover relationships via linear and non-linear multivariate analysis of data across many stations.

It was initially intended that this research would be a multivariate analysis across a distributed gauging system. It was quickly discovered, however, that the necessary “building blocks” for a study of that kind were missing. Therefore, this research established methodologies that could be used to prepare data for studying the effects that upstream inputs can have on the downstream body in white (BIW).

Four methodologies were presented: tracking parts, removing outliers, identifying presentation error, and evaluating the BIW results. These methodologies are essential in conducting variation reduction over a distributed gauging system. While these methodologies have been applied to data from an automotive shop, they should be applicable in other manufacturing settings as well.

Tracking Parts

For a small cost, the GM assembly plant in this research has tapped into an ability that few plants can emulate: tracking of all major sub-assemblies to the BIW. The solution and methodology for tracking parts, an information system approach, are very plant specific. Other plants, even within General Motors, would have difficulties directly adopting the methodology implemented in this body shop. Support and involvement of plant personnel were a critical component needed to ensure the integrity of the tracking.

Removing Outliers

When an analysis indicates something is wrong, it is probably true. It may be, though, that the problem is with the measurements! Analyses that are not robust to outliers are only as good as the data provided to them, therefore outliers must be screened and removed.

Team involvement in root cause identification and solutions remained the best way to reduce outliers. Prevention of all outliers is impossible, however, and some outlier screening will remain in statistical routines. Multivariate outlier screening methods appear to have an advantage over univariate screening particularly when the data variables may be related.

Outlier removal on the data sets from the Optical Coordinate Measuring Machines drastically reduced the number of pure records. Missing measurements also played a role in the data set reduction. A methodology for replacing outlier and missing measurements would be a great area for future research.

Identifying Presentation Error

A significant amount of variation in the data was caused by presentation error, error resulting from play in the way the part is located when measured. A new, non-iterative methodology for removing and identifying presentation error was developed because no current algorithms could simultaneously address all of the limitations imposed by the data from the assembly plant. These limitations included the limited number of data points, the three dimensional translations and rotations, and that few of the data points had measurements in all three dimensions.

There is some progress and improvements that can be made to the presentation error methodology. The methodology applied was not robust to missing data or outliers. For that reason, the data was screened for these conditions prior to running through the algorithm. When the number of data points is large, such as with the BIW measurements, improvements in robustness to outliers should be possible by using more robust linear regression techniques.

Evaluating the BIW Results

Providing a mechanism or methodology for interpreting the data is as importantant as providing the good, clean data. A proper interpretation must look at the relationships between the data variables. Capturing these relationships is important because it helps in the understanding of how the BIW is building. Being able to watch key relationships helps simplify control and variation reduction in the body shop. Relationships help in comparing different zones on the BIW as well as in comparing different BIWs.

Relationships also help during some analyses by reducing the dimensionality of the data set. This reduction is useful in the body shop because there are typically more variables measured in the downstream BIW gauge than in the upstream gauges. Relationships reduce the complexity of the analysis while retaining much of the information stored in the data.

It is troubling to see that so few relationships are being used in the body shop. There is much to be gained by monitoring significant relationships. More troubling is that each time advanced relationships are used, the "wheel" seems to be re-invented. This GM body shop should provide continued evidence over the next year on how relationships implemented in-line can help in proactive control and variation reduction in the body shop.

References

Arun, K.S., T.S. Huang and S.D. Blostein. "Least-Squares Fitting of Two 3-D Point Sets." IEEE Transactions on Pattern Analysis and Machine Intelligence. vol. PAMI-9, No. 5, September 1987: 698-700.

Baron, Jay S. "Dimensional Analysis and Process Control of Body-In-White Processes." Diss. University of Michigan, 1992.

Choi, Woncheoi and Thomas Kurfess. "Data Localization Algorithms and Minimum Zone Evaluation for Automated Inspection." International Mechanical Engineering Congress and Exposition. DE-Vol. 74, 1994: 39-46.

Clinton, James. "Body Shop - Variation Reduction Cartrac Problem Solving Team - Problem Solving Sheets." General Motors, December 9, 1994.

EDS Service Proposal. "Data Traceability and Continuoius Improvement Indicator (CII) Support." document 190182, December 13, 1994.

Dewar, Robert. "Sub-pixel Measurement with Structured Light Sensors." TechLink Associates. Novi, MI., 1994.

Gretz, Bruce. Personal interview. March 17, 1995.

GM brochure. Passage to the Future. BOC-1-5M-1286, General Motors, 1986.

Harvey, Dennis. "Front End Sheet Metal Fit Study." General Motors, March 26, 1982.

Hoaglin, D. C., F. Mosteller, and J. W Tukey. Understanding Robust and Exploratory Data Analysis. New York: Wiley, 1983.

Horn, Berthold K. P., Hugh M. Hilden, and Shahriar Negahdaripour. "Closed-form solution of absolute orientation using orthonormal matrices." J. Opt. Soc. Am A, vol. 5, No. 7, July 1988: 1127-1135.

James, Gregory A. "Process Monitoring Methodologies for Sheet Metal Assembly Operations." Diss. Massachusetts Institute of Technology, 1993.

Jinkerson, R. A., S. L. Abrams, L. Bardis, C. Chryssostomidis, A. Clement, N. Patrikalakis, and F. Wolter. "Inspection and Feature Extraction of Marine Propellers." Journal of Ship Production. vol. 9, no. 2, May 1993: 88-106.

McGill, R., J. W. Tukey, and W. A. Larsen. "Variations of box plots." The American Statistician. vol. 32, 1978: 12-16.

Meyer, W. Weston. "2D results." General Motors Technical Center. 1994.

Meyer, W. Weston. "Methodology Comparison." General Motors Technical Center. January 14, 1995.

Perceptron Algorithm Manual. PN #009-0073, June 1993.

Perceptron DSP 1000 System Manual. PN #009-0079, January 1994.

Perceptron Visual Fixturing. PN #009-0076, October 1993.

Pizzimenti, David J. "Perceptron, Inc." Nomura Research Institue America, Inc. New York, N.Y., 1994: 1-16.

Pool, Kirby V. "Product/Process Improvement through Performance Measurements and Modeling: a Case Study in Aerospace Welding." Diss. Massachusetts Institute of Technology, 1991.

Schonemann, Peter H. "A Generalized Solution of the Orthogonal Procrustes Problem." Psychometrika. vol. 31, no. 1, March 1966: 1-10.

Sharma, V., R. E. York, R. E. Welsch, S. D. Eppinger. "Non-iterative data localization algorithm for partial and missing measurements on a body." 1995. Under preparation. For correspondence, contact Prof. Roy E. Welsch, MIT Sloan School of Management, 30 Wadsworth St., E53-383, Cambridge, MA 02142.

Tukey, J. W. " Data-based graphics: visual display in the decades to come." Statistical Science. vol. 5, 1990: 327-339.

Velleman, P. F., and D.C. Hoaglin. Applications, Basics, and Computing of Exploratory Data Analysis. Boston: Duxbury, 1981.

Appendix A - Tracking Proposal Summary

Data Traceability and Continuous Improvement Indicator (CII) Support

Summary of EDS Service Proposal 190182
Presented to
GENERAL MOTORS
December 13, 1994

Current Environment
Body Shop and NAO personnel are manually correlating data from eight stand-alone Optical Coordinate Measuring Machines (OCMM). For purposes of Body-In-White (BIW) process control, the collection and analysis of this information allow for rapid investigation and identification of Body Shop variation.

Continuous Improvement Indicator (CII) Software is one example of analysis presently used by the Body Shop.

Proposed Solution
EDS will research and develop a system to allow data traceability among (four) sideframes, Body-In-White, and cartrac pallet OCMMs. EDS will provide a centralized personal computer which will be used as a host for GM's Continuous Improvement Software, data downloading, data analysis, and variation reduction. EDS will make CII software and personal computer configuration changes to accept additional data. EDS will also provide FTP software to allow for remote system access.

EDS will also provide CII software support services following successful installation of hardware and software infrastructure. This software support includes: Software maintenance, Software enhancements to reflect model changes, Chart and Report Changes, On-site support 1st shift, and Technical support by pager for 2nd shift.

Benefits
- On-site Continuous Improvement Indicator software support.
- Remote/LAN access to data.
- Traceability of data and therefore traceability of work-in-process.
- On-site systems support for software and hardware.
- Enabling the plant to re-deploy existing resources.
- A centralized data collection point.

Appendix B - BIW Ruler Examples

Relationship 1 is Front width of roof

RELATIONSHIP MEASUREMENT DATA
ROOF RELATION 1
ISN 5086152 TO 5086168 12/13/94 15:40 TO 12/13/94 15:56

16:19:5
12/13/9
SAMPLE OF 1

MEASUREMENTS (MM)

JSN	-/+ I/O
5086168	-0.22
5086166	-0.23
5086165	-0.15
5086163	-0.22
5086161	-0.21
5086159	-0.22
5086158	-0.21
5086156	-0.22
5086154	-0.16
5086152	-0.17

example:

-.22% means too narrow

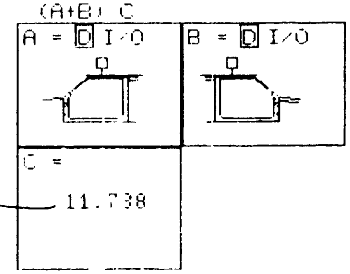
(-.22 x 11.738) = -2.58mm

∴
The front is building -.22% smaller than design. That is equivalent to -2.58mm

MEASUREMENTS (MM)

JSN	- + I/O

RELATIONSHIP :
(A+B) C



1

OPER: I
BODY DAT
NON-CRIT
PALLET # = AN
SHIFTS = AN
QUALITIES = AN
UNFIXTURED

Relationship 2 is Back width of roof

RELATIONSHIP MEASUREMENT DATA　　　16:20:01
ROOF　　RELATION 2　　　12/13/94
JSN 5086152 TO 5086168　12/13/94 15:40 TO 12/13/94 15:56　　SAMPLE OF　10

MEASUREMENTS (MM)

JSN	-/+ I/O
5086168	-0.23
5086166	-0.25
5086165	-0.25
5086163	-0.25
5086161	-0.23
5086159	-0.25
5086158	-0.24
5086156	-0.23
5086154	-0.21
5086152	-0.23

Example

-0,23% means too narrow

(-0.23 × 11.528) = -2.65mm

The back of the roof is Too narrow by -0.23% or -2.65mm.

MEASUREMENTS (MM)

JSN	-/+ I/O

RELATIONSHIP :
(A+B)/C

A = [] I/O　　B = [] I/O

C =
11.528

OPER: 0
BODY DATA
NON-CRIT.
PALLET # = ANY
SHIFTS = ANY
QUALITIES = ANY
UNFIXTURED

2

relationship 3 is A or V build of roof

RELATIONSHIP MEASUREMENT DATA
ROOF RELATION 3
JSN 5086152 TO 5086168 12/13/94 15:40 TO 12/13/94 15:56

16:20:08
12/13/94
SAMPLE OF 10

MEASUREMENTS (MM)

JSN	-/+ I/O
5086168	0.00
5086166	0.01
5086165	0.10
5086163	0.02
5086161	0.01
5086159	0.03
5086158	0.02
5086156	0.00
5086154	0.05
5086152	0.06

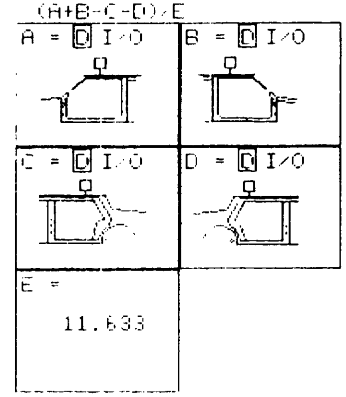
a positive % means wider in front than back.

ie. roof is building like

MEASUREMENTS (MM)

JSN	-/+ I/O

RELATIONSHIP :
(A+B-C-D)/E



OPER: 0
BODY DATA
NON-CRIT.
PALLET # = ANY
SHIFTS = ANY
QUALITIES = ANY
UNFIXTURED

3

relationship 4 is skewness

RELATIONSHIP MEASUREMENT DATA
ROOF RELATION 4
JSN 5086152 TO 5086168 12/13/94 15:40 TO 12/13/94 15:56

16:20:16
12/13/94
SAMPLE OF 10

MEASUREMENTS (MM)

JSN	-/+ I/O
5086168	0.10
5086166	0.07
5086165	0.07
5086163	0.05
5086161	0.09
5086159	0.05
5086158	0.00
5086156	0.07
5086154	0.05
5086152	0.06

positive % means front is building to the left.

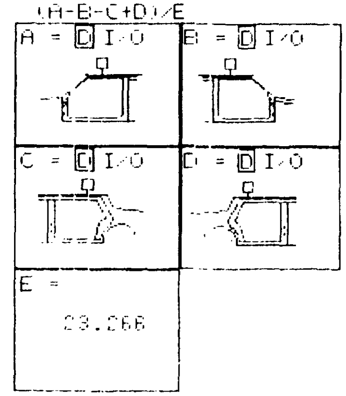
roof

front

MEASUREMENTS (MM)

JSN	- + I/O

RELATIONSHIP :
(A-B-C+D)/E



OPER: 0
BODY DATA
NON-CRIT.
PALLET # = ANY
SHIFTS = ANY
QUALITIES = ANY
UNFIXTURED

4